DOCUMENT RESUME

ED 274 141 FL 015 901

AUTHOR Derbyshire, Desmond C., Ed.
TITLE Work Papers of the Summer Institute of Linguistics, University of North Dakota Session, Volume 30, 1986.
INSTITUTION Summer Inst. of Linguistics, Grand Forks, N. Dak.
PUB DATE 86
NOTE 149p.; For individual papers, see FL 015 902-909.
PUB TYPE Collected Works - General (020) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS Eskimo Aleut Languages; Foreign Countries; Form Classes (Languages); *Grammar; Language Planning; Lexicology; *Linguistics; *Structural Analysis (Linguistics); Uncommonly Taught Languages
IDENTIFIERS Kanasi; Kope; Lakota (Language); Majang; Mazatec; Orthography; Piraha; Zambales Mountains

ABSTRACT

Working papers resulting from the 1986 University of North Dakota Summer Institute of Linguistics include: "Orthographic Reform in Kope" (John M. Clifton); "Ternarity and Obligatory Branching in Piraha" (Daniel Everett); "Reduplication in Majang" (Pete Unseth); "Indirect Objects and Incorporation in Mazatec" (Daniel Agee, Stephen Marlett); "Grammatical Relations in Eskimo: A Response to Kalamar" (Richard J. Nivens); "A Relational Grammar Approach to Verb Agreement in Lakota" (Gray Plunkett, Michael McKeever); "Kanasi: A Brief Grammar Sketch" (Ronald W. Pappenhagen); and "The Languages of the Zambales Mountains: A Philippine Lexicostatistic Study" (John Wimbish). (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED2741141

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Derbyshire

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

WORK PAPERS

VOLUME XXX

1986

UND

SUMMER INSTITUTE OF LINGUISTICS
UNIVERSITY OF NORTH DAKOTA SESSION

BEST COPY AVAILABLE

FL 015 901

1986

WORK PAPERS

of the

Summer Institute of Linguistics

University of North Dakota
Session

Volume 30

Editor: Desmond C. Derbyshire

Summer Institute of Linguistics

These are working papers and should not be cited without referring to their preliminary nature.

Orders and correspondence should be sent to:

SIL/UND Work Papers
c/o International Linguistic Center-Bookroom
7500 West Camp Wisdom Road
Dallas, TX 75236

PREFACE

Most of the papers in this volume present some provisional results of research being undertaken by members of the Summer Institute of Linguistics who are associated with the SIL's summer program at the University of North Dakota, either as teachers or research students.

My thanks go to Diana Weber for handling the composition and production of the volume and to John Clifton, Betty Brown, Steve Marlett, Tom Payne, Steve Quakenbush and Chuck Speck for their editing help.

D.C.D.

TABLE OF CONTENTS

Preface iii

Orthographic Reform in Kope
John M. Clifton 1

Ternarity and Obligatory Branching in Pirahã
Daniel Everett 13

Reduplication in Majang
Pete Unseth 43

Indirect Objects and Incorporation in Mazatec
Daniel Agee and Stephen Marlett 59

Grammatical Relations in Eskimo: A response to Kalamár
Richard J. Nivens 77

A Relational Grammar Approach to Verb Agreement in Lakota
Gray Plunkett and Michael McKeever 89

Kanasi: A Brief Grammar Sketch
Ronald W. Pappenhagen 107

The Languages of the Zambales Mountains: A Philippine Lexicostatistic Study
John Wimbish 133

ORTHOGRAPHIC REFORM IN KOPE

John M. Clifton

1 Introduction
2 Previous orthographic tradition
3 Nasal / fricative variation
4 Vowel clusters vs. semivowels
5 Long vowels and glottal stop
6 Word breaks
7 Conclusion
Notes
References

1 Introduction

Language planners trying to formulate a workable orthography in countries like Papua New Guinea often find at least a core of speakers who have strong feelings about how their language should look on paper. Although some of these feelings may well come from intuitive, language-internal sources, many come from sociological sources such as early attempts by foreign missionaries to write the language, or through knowledge of written trade or national languages.

These feelings often lead to inadequate or, when they arise from different sources, conflicting orthographies. Even within a single orthographic 'tradition' there are generally inconsistencies between writers in spelling and word breaks. Further problems frequently arise from dialect differences or complex morphophonemic systems.

These problems make orthographic reform desirable in many extant systems. Any such reform effort will, however, probably encounter resistance due to the sociological pressures. In some areas of Papua New Guinea the current orthographic tradition has been in place for 100 or more years. Although the practical considerations are not as great as, for example, in European languages, those literate in the current orthography as well as churches or other institutions with literature in the previous orthography have a vested interest in maintaining the status quo.

In the light of these problems, some language workers have decided it is futile to attempt orthographic reform regardless of what problems may exist due to a mismatch between the orthography and phonological system. Others

simply ask people how they would like to write particular words, and take the majority opinion at face value. A third approach is to base the orthographic decisions on the phonological analysis and insist on this as the 'scientific' system.

It is my contention that while reform may be difficult and time-consuming, most people are open to reform if they can see the advantages from such a change outweigh the disadvantages in reorienting current literates and possibly reprinting certain books (although it is probably likely the two systems could coexist with regard to previously published literature). The first option noted above does not even raise this as a possiblity.

While it is necessary to involve the local people in the decision making, always following majority decisions may well lead to problems, as those responding will probably not recognize the full range of options available to them or the implications of various decisions. It seems the role of the language planner must include explaining such issues to those making orthographic decisions.

Finally, in addition to the sociological factors ignored by an approach basing orthographic decisions solely on the phonological analysis, it does not allow the analyst to take advantage of insights into the phonological system which might be afforded by reactions to orthographic decisions. In many cases such reactions indicate problems in the phonological analysis.

In this paper, we will examine how local attitudes played a role in orthographic reform in Kope. In section 2 the previous orthographic tradition is outlined. Then in sections 3 and 4 we examine areas of apparent concern where local feelings pointed to new understanding of the phonological system and no change was made. Finally, in sections 5 and 6 we examine areas in which a genuine case for reform could be made and changes resulted.

2 Previous orthographic tradition

Kope, Urama, Gibaio, Maipu'a, and Arigibi are the five dialects that make up the easternmost language in the Kiwai language family.[1] Missionaries from the London Missionary Society (LMS) began work in two Kiwaian languages, Southern Kiwai and Kerewo, in the late 1800's and early 1900's, respectively. They also worked in several unrelated languages to the east of Kope.

The only materials published in Kope or related dialects were a number of choruses in Kope and Urama included in a hymn book (Gido Buka), and a Kope translation of Luke 2:1-20. Much more vernacular literacy work was attempted by the LMS in Kerewo, however. Gido Buka includes a larger number of Kerewo choruses and translated hymns, and a Kerewo translation of the Gospels and Acts was published. The LMS also set up schools taught in Kerewo. As Kope and Kerewo are very closely related linguistically and socially, Kope speakers attended Kerewo schools and learned to read Kerewo. Thus, in spite of the lack of published Kope literature, many Kope speakers wrote letters and posted 'public service' announcements in Kope.

The consonant system throughout the Kiwaian languages is very similar and, with the exception of the glottal stop, does not make use of nonEnglish contrasts. In Kope, the following English consonants were used: **p t k b d g h v m n r**[2]. Glottal stop was symbolized by **'** in all the languages in which the LMS worked. In addition, **s** was often used in loan words with this sound.[3]

The vowel system used the five vowels, **a e i o u**, with the values [a ɛ i o u]. No semivowels were used; instead, vowel clusters were written whether or not there was a syllable break. Examples are given in (1-2).

(1)	[maᵘ.a][4]	**maua**	'sibling'
	[meⁱveⁱ]	**meimei**	'pineapple'
	[mihaᵉ]	**mihae**	'rain'
	[ⁱamagaᵘri]	**iamagauri**	'to jump'
(2)	[me.a]	**mea**	'good'
	[idi.o]	**idio**	'to drink'
	[to.ɛ]	**toe**	'frightened'

3 Nasal / fricative variation

One orthographic problem involved near-allophonic variation between [m] and [v][5], and [n] and [r]. In most Kiwaian languages these are separate phonemes, but in Kope and related dialects they have almost completely collapsed with the nasals occurring word initially and the fricatives[6] occurring intervocalically. Examples are given in (3-4).

(3)	[mahoroo]	**mahoroo**	'bandicoot'
	[mɛveraa]	**meneraa**	'flying marsupial'
	[mɛviho]	**meniho**	'bad'
	[ɛvapua]	**enapua**	'in-law'

(4)	[naara]	**naara**	'things'
	[nɛhɛbo]	**nehebo**	'tree sp'
	[mario]	**mario**	'bird sp'

This variation at the two points of articulation appears to be parallel and would thus theoretically be the result of a single rule. This further implies they should be treated the same orthographically, preferably writing only one allophone in each case. Common orthographic practice as seen in (3-4), however, treated them differently. While **m** was generally used in all positions, **n** was restricted to initial position while **r** was used intervocalically.

Closer phonological examination shows that in fact the two are not totally parallel. When most words containing intervocalic /m/ are pronounced carefully, they exhibit [m], not [v]. Many speakers emphatically deny any of the words in (3) are ever pronounced with [v], although the fricative is clear to nonKopes. On the other hand, no words are ever pronounced with intervocalic [n], even in the most careful speech. Although speakers can be made aware that [n] never occurs intervocalically and that [r] seldom occurs word initially and so no confusion results from writing both as **n**, they insist [n] and [r] are distinct.

Another difference between the two points of articulation is reflected in exceptions. The three words in (5) are always pronounced with an intervocalic [v].

(5)	[hivio]	**hivio**	'sun'
	[ovɛ]	**ove**	'moon'
	[orovidio]	**orovidio**	'to hear'

The exceptionality of these forms is not that they have the wrong allophone, but that they are not pronounced with [m] even in the most careful speech.

The exceptions to the alveolar variation are given in (6).

(6)	[rautu]	**rautu**	'with'
	[rio]	**rio**	'desire'
	[ro]	**ro**	'you (sg)'

These forms apparently violate the allophonic distribution, since they have word initial [r]. It is interesting, though, that while all three forms in (5) are nouns, none of the forms in (6) are. This may account for the [r] in [rautu], as it is a postposition and would generally be unstressed.

[rio] also tends to be subordinate to the preceding noun or pronoun as in (7).

(7) ['mo rioka 'duu]
1sg desire food
'My desire is food.'

A similar analysis cannot be provided for [ro], however, especially as the pronouns [nu] '3sg', [ni?o] '2pl', and [nii] '3pl' all begin with [n].

Although no explanation has been found yet to account for these differences in behvaior between the bilabial and alveolar points of articulation, the differences are clear. At an orthography conference called primarily to discuss long vowels (see Sect. 4), the participants decided to follow the prevailing convention of writing all instances of [m] or [v] as **m** except for the three forms in (5), but to differentiate between [n] and [r], writing **r** intervocalically and initially in the forms in (6) and **n** initially in all other forms.

The use of both **n** and **r** did raise problems for the representation of the verb prefix indicating first person involvement in the sentence. Examples are given in (8). (S = 1st person involvement, T = tense)

(8) a. [mo n-o'u-baka do'ou]
1sg S-come-T today
'I came today.'

b. [mo pi-r-o'u goroi purai]
1pl T-S-come before week
'I came last week.'

In (8a) the prefix is word initial and so takes the form [n], while in (8b) it is intervocalic and so takes the form [-r-]. The decision made at the orthography meeting in this case was to maintain morpheme identity by writing **r** in all cases.[7]

4 Vowel clusters vs. semivowels

A second apparent orthographic problem was the lack of the semivowels **w** and **y**. It seemed this lack would make it difficult for readers to identify syllables and would thus limit the use of syllables in literacy. Second, there appeared to be minimal pairs which the current system could not differentiate. One such pair is given in (9).

(9) a. [u.ɛ] **ue** 'sugar cane'
b. [ᵘɛ] **ue** 'arrow type'

There was general resistance, however, to the idea of including **w** and **y** in the orthographic inventory.

Some of this resistance was due to the fact that initial /**m**/ in Kope is cognate with /**w**/ in the Gibaio dialect, as seen in (10).

(10)	Kope	Gibaio	
	/**m**ade/	/**w**ade/	'word, language'
	/**m**apo/	/**w**apo/	'tail'

Thus, to write initial **w** in Kope made it 'look like Gibaio'. In addition, it became apparent listening to people read Kope that while the vowel clusters might be problematic for us, they did not seem to be problematic for native speakers.

Further analysis indicated a phonological basis for this resistance to distinguishing [u] and [w], and [i] and [y] orthographically. In most instances, the syllabification of a vowel cluster as one or two syllables is predictable from the following principle.

Syllabification of Vowel Clusters (SVC)

a. If a given vowel is articulatorily lower than the following vowel, they are syllabified as one syllable, with the higher vowel becoming nonsyllabic.
b. A word initial high vowel becomes nonsyllabic before another vowel, with the two vowels being syllabified as a single syllable.
c. Otherwise, consecutive vowels are syllabified as separate syllables.

SVC accounts for all the forms in (1-2) above, repeated here as (11-12).

(11) a. [mɑᵘ.ɑ] **mɑua** 'sibling'
b. [meⁱveⁱ] **meimei** 'pineapple'
c. [mihɑᵉ] **mihae** 'rain'
d. [ⁱɑmɑgɑᵘri] **iamagauri** 'to jump'

(12) a. [me.ɑ] **mea** 'good'
b. [idi.o] **idio** 'to drink'
c. [to.ɛ] **toe** 'frightened'

In the case of (11a) SVC(a) predicts **u** will syllabify with the preceding **a**, and SVC(c) predicts there will be a

syllable break between it and the following **a**. The desyllabification of the initial **i** in (11d) follows from SVC(b). The other cases of desyllabification in (11) follow from SVC(a).

SVC(b) also accounts for the desyllabification of the initial **u** in (9b). It does not, however, account for the contrast in (9). This problem will be resolved in section 5. In the meantime, it is significant that SVC accounts for the vast majority of syllable breaks.

5 Long vowels and glottal stop

Undoubtedly the most difficult orthographic problem area was that of long vowels. Kope clearly distinguishes vowel length and/or stress or tone. Representative contrasts are given in (13-16).

(13)	['abeɑ]	**abea**	'father'
	[ɑ'beeɑ]	**abea**	'sago squeezing bag'
(14)	['niivo]	**nimo**	'we'
	[ni'vo]	**nimo**	'lice'
(15)	['oobo]	**obo**	'woman'
	[o'bo]	**obo**	'water'
(16)	[tu'tuu]	**tutu**	'long'
	['tutu]	**tutu**	'handle'

While a few Kope speakers describe the contrasts in the above words in terms of the 'melody' (stress and tone seem tied together), most describe it in terms of 'stretch'. As indicated in (13-16), the difference was not indicated in the orthography, however, so readers had to rely on context to differentiate minimal pairs.

From the beginning we encountered resistance in our attempts to indicate vowel length by the commonly used convention of geminate vowels. Some of the verbalized resistance was based on a dislike of the looks of so many geminate vowels and and a feeling that there were few potential ambiguities. A further problem with the symbolization of long vowels as geminates relates to the representation of glottal stop. As was mentioned above, glottal stop was written as '. This ' was frequently omitted when writing, however, especially when the two vowels were identical. Typical spellings are given in (17-18).

(17) a.	[oʔu]	**o'u**	'to come'

	b.	[ɛvɛʔio]	**eme'io**	'to burn garden area'
	c.	[oruʔo]	**oru'o**	'to wash oneself'
(18)	a.	[goʔotɑ]	**goota**	'coconut'
	b.	[ɛvɑʔɑi]	**emaai**[9]	'to give'

The widespread use of ' in (17) where the vowels are dissimilar is due to the fact that both vowels in a sequence of dissimilar vowels are generally written. Thus, (17c) would be identical with (19) if ' were omitted.

(19)	[oruo]	**oruo**	'to climb down'

It is obvious that [oo] and [oʔo], for example, could not both be spelled **oo**. If geminates were used for long vowels, ' would have to be written between identical as well as dissimilar vowels. The forms in (18) would then be spelled **go'ota** and **ema'ai**, respectively. This was, indeed, how some people spelled these words. Although most people did not object to writing ' between identical vowels, they still felt that the habit of reading geminate vowels as separated by a glottal stop was too strong to overcome, and that spelling [oobo] 'woman' as **oobo** would likely lead to a pronunciation of [oʔobo].

While the actual representation of long vowels was problematic, work on the lexicon turned up an increasing number of minimal pairs contrasting in vowel length. Three options for indicating long vowels without using geminates were: a macron over the vowel (**ōbo**), a hyphen after the vowel (**o-bo**), or a colon after the vowel (**o:bo**). When these possibilities were presented to two Kopes involved in Bible translation, both disliked the colon as this would involve 'punctuation in the middle of a word'. After some discussion, they decided to try a macron, based on their familiarity with the dictionary. This option was rejected, however, when they came to realize this meant both backspacing and rolling the platen down and up when using a standard typewriter.[9]

Later, participants in two separate literature production workshops decided to try the hyphen to indicate long vowels. (The colon was once again rejected as being punctuation.) This also proved problematic because the writers frequently broke words between lines due to their length. This led to confusion as to whether a hyphen at the end of a line represented the continuation of a word or a long vowel.

As a result of these practical considerations, we were tempted to underdifferentiate and continue writing both long and short vowels as single vowels. This seemed reasonable, in spite of the apparently high semantic load carried by length differences in isolation forms, as there were very few minimal pairs which could not be disambiguated by context. In spite of the contextual clues, however, one of the Kope translators noticed that when people read his translation to check for naturalness, they consistently had problems with certain words containing long vowels. When he indicated the long vowels with geminates these problems were almost completely eliminated.

At this point a general meeting of a number of Kope speaking pastors, teachers, medical orderlies, village counselors, and translators was called to discuss orthography matters. The main issue was the representation of long vowels. As we discussed the various options available the general concensus to try writing long vowels with geminates and then see if older pastors could read the new orthography without major problems. The use of geminates in the recent publication of Christmas story seems to have been well accepted.

The decision to indicate vowel length has also helped distinguish between (9a) and (9b), repeated here as (20a-b).

(20)	a.	[u.ɛ]	**ue**	'sugar cane'
	b.	[uɛ]	**ue**	'arrow type'

In discussing these forms with various people we found that most felt this distinction was one of length. Once length is represented, 'sugar cane' is **uue** as opposed to 'arrow type' which is still **ue**. This gives further evidence for the need to differentiate long and short vowels in the orthography.

6 Word Breaks

A final orthographic problem was the inconsistent indication of word breaks. Three sets of forms provided most of the problems. The first two were reduplicated forms, illustrated in (21), and verbal inflections, illustrated in (22).

(21)	a.	**himia himia** or **himiahimia**	'different'
	b.	**hiiro hiiro** or **hiirohiiro**	'plenty'
	c.	**ha'ima ha'ima** or **ha'imaha'ima**	'very tired'
	d.	**ora ora** or **ora'ora**	'red'
	e.	**otohiti otohiti** or **otohiti'otohiti**	'wrapped repeatedly'

(22) a. **odau kaida** or **odaukaida** 'they (2) are going'
b. **emidio kaumo** or **emidiokaumo** 'they are staying'

Phonologically, all these examples are single words. Thus, writing them as two words led to problems in reading with normal intonation. Writing them as one word, on the other hand, resulted in long words many readers found difficult to decode. There was also a problem as to whether to write the glottal stop in reduplicated forms when the base form began with a vowel as in (21d-e). The glottal stop is predictable in careful speech in this environment although it is frequently omitted in fast speech.

Once the decision was made not to use the hyphen for long vowels, the group mentioned above decided to use it here. This decision was made on the basis of the use of the hyphen in English to join closely related words. The group felt forms like 'close-up' and 'half-baked' especially paralleled reduplicated forms in Kope. This use of the hyphen in **himia-himia**, **otohiti-otohiti**, and **odau-kaido** resulted in single words which could be more easily decoded. In addition, it eliminated the need to indicate glottal stop in reduplicated forms.

The other problematic forms involved the ergative marker **-ro** which is attached to the final word in the noun phrase. This suffix was generally written separately from proper nouns, but attached to other words. Examples are given in (23).

(23) Ga'ia ro peidai.
Ga'ia SBJ took
'Ga'ia took it.'

Niiro nai pihoumo.
3pl-SBJ fish ate
'They ate fish.'

Nu ekeiro omihia.
3sg branch-SBJ hit
'The branch hit him.'

Bomo gemairo piho.
pig big-SBJ ate
'The big pig ate it.'

These conventions were not universally followed, however. The decision to use hyphens in (21) and (22) was therefore

extended to apply to the ergative, resulting in **Ga'ia-ro, nii-ro, ekei-ro** and **gemai-ro**.

7 Conclusion

Orthographic reform is a complex issue. While many of the problems arise from sociological factors, others arise from the interaction of the phonological analysis and the proposed reforms. On the one hand, resistance to such reform may well point up problems in the underlying phonological analysis. On the other hand, a good analysis will often provide the basis on which people can make informed decisions regarding orthography.

Notes

1. Wurm (1973) calls this language Northeast Kiwai, since there is no local name for the overall language. He includes four dialects in Northeast Kiwai, Gope (Kope), Gibaio, Urama, and Baravi (Maipu'a). I include Arigibi, which Wurm lists as a separate language, as a fifth dialect.

2. Bold face type is used for orthographic representations.

3. Kope /h/ is cognate with /s/ in neighboring languages and so **s** is present in these nearby orthographies.

4. This could also be broken as [mɑ.ᵘɑ]. The rationale for breaking it as above will be outlined in section 5.

5. [v] fluctuates with [ƀ] and even [b] in some speakers.

6. [r] is the usual alveolar fricative in most Papuan languages.

7. The decision to use **r** instead of **n** was made on the basis that while [r] sometimes occurs initially, [n] never occurs intervocalically.

8. People often write the forms in (18) with only single vowels, but this is generally considered to be incorrect.

9. This would not be a problem on a typewriter with a macron on a dead key. There are a number of standard typewriters scattered throughout the language group, however, which would be used regularly.

References

Wurm, S. 1973. The Kiwaian languages. In The Linguistic Situation in the Gulf District and Adjacent Areas, Papua New Guinea, ed. K. Franklin. Pacific Linguistics, series C, no. 26. Canberra: Linguistic Circle.

TERNARITY AND OBLIGATORY BRANCHING IN PIRAHÃ*

Daniel Everett

1 Outline
2 The basic facts
3 Two failed analyses
  3.1 The grid
  3.2 Binary trees
4 Ternarity and obligatory branching
  4.1 Amphibrachs
  4.2 Stress in the lexicon
  4.3 Summary
5 Morphemic fusion
6 Extrametricality
7 Conclusion
Notes
References

1 Introduction

A central concern of multilinear phonology, as developed by Liberman and Prince (1977); Hayes (1981); Prince (1983); Hammond (1984); Halle and Vergnaud (forthcoming); and others, has been the characterization of the major components of prosodic structure: **hierarchy, directionality**, and **quantity**. Hierarchy is the organization of prosodic units into phonological constituents, e.g. syllables and feet. Directionality refers to the general orientation, left-to-right, of phonological process at a given structural level. Quantity may be understood as relative structural complexity, expressed in terms of particular notational devices (such as branching or grid complexity) or by less theoretical vocabulary, such as syllable weight.

In this paper it is argued that the analysis of stress placement in Pirahã offers important new evidence on the proper characterization of prosodic organization as well as the notational devices necessary to express this organization[1]. Specifically, I will show that Pirahã stress placement is intractable to either a **grid** account (Prince 1983) or a binary branching analysis but that an insightful perspective on a variety of stress related phenomena is available if we assume that Pirahã has ternary, **obligatory branching** feet. Evidence is presented from the assignment of stress in morphologically simple words, lexical

compounds, words with **extrametrical** suffixes, and transmorphemic stress shifts and clash resolution.

2 The basic facts

The rule of stress placement in Pirahã may be stated as in (1) (see also Everett and Everett (1984) and Everett (1985; 1986, Sect. 22)):

(1) Pirahã stress placement (initial statement):
Stress the rightmost token of the heaviest syllable type in the word.

Syllable weight relations are expressed (2): (where **C** = voiceless consonant; **G** = voiced consonant; . = 'is heavier than')

(2) CVV>GVV>VV>CV>GV

Rule (1) correctly predicts stress placement in the following examples (bold = stress; . = syllable boundary).

(3) Rightmost Token

a.	paó. hoa. **hai**	'anaconda'
b.	ba. hói. ga. **toi**	'domesticated pig'
c.	bai. tói. **sái**	'wildcat'

(4) CVV vs GVV

a.	bii. **sai**	'red'
b.	**kai**. bai	'species of monkey'

(5) CVV vs VV

a.	xi. **sii**. ho. ái	'liquid fuel'
b.	ko. so. ii. gai. **tái**	'eyebrow'

(6) CVV vs CV

a.	**xoo**. gi. ái.	'proper namé
	xo. gi. **ái**	'big'
b.	**tai**. si	'axe'
c.	si. **tai**	'feather'

(7) CVV vs GV

a. **xaa**. gí	'coati mundi'
b. ga. **hió**	'airplane'

(8) GVV vs VV

a. **gao**. ii	'proper name'
b. ho. aa. **gai**	'species of fruit'

(9) GVV vs CV

a. xi. **gaó**	'back (body part and locational)'
b. **gái**. hi	'that'

(10) GVV vs GV

a. bó. **gai**	'breast'
b. **gaa**. ba	'next'

(11) VV vs CV

a. pí. **ai**	'also'
b. ho. **aí**. pi	'proper name'

(12) VV vs GV

a. xa. **aí**. bi	'thin'
b. si. ga. **ii**	'fins'

(13) CV vs GV

a. **xa**. ba gi	'toucan'
b. xa. ba **pa**	'proper name'

An important theoretical implication of Pirahã stress has already been pointed out in Everett and Everett (1984) and Davis (1985), namely, that the stress placement rule crucially refers to the syllable onset. Consequently, contrary to many current works (see especially Hyman 1985), stress placement may be based on the **syllable projections** rather than merely **rime** or **nucleus projections**.

I want to focus here, however, on a different but equally important facet of Pirahã stress - the restriction of stress to the final three syllables, as illustrated in (14)-(18):

(14)	pia. hao. gi. so. **aí**. pi	'cooking banana'
(15)	poo. **gái**. hi. aí	'banana'
(16)	kao. **ái**. bo. gi	'jungle spirit'
(17)	ka. pii. ga. ií. to. **ii**	'pencil'

(18) poo. so. gi. **ái** 'species of fish'

In each of the examples in (14)-(18), stress can only be placed on one of the final three syllables, in spite of the fact that a heavier syllable type (CVV) occurs to the left of the antepenult.

Superficially, this restriction does not look too unusual. It is certainly easy to state in prose. Also, it is not perhaps immediately apparent that Pirahã differs radically from better known languages, such as Spanish. As Harris (1983) shows, Spanish stress is limited to the final three syllables, interacts with syllable weight, and may appear on either the antepenult, the penult, or the ultima. On closer examination, however, the way in which the domain of stress placement is determined in Spanish (and other similar cases I am aware of) is drastically different than in Pirahã. According to Harris' (1983) analysis, noun stress in Spanish is penultimate, except in words like **papá** 'dad' which have no gender suffix (what Harris (1983, 91ff) calls the **Terminal Element**). Apparent counterexamples, such as **nómada** 'nomad' (cf. **pomáda** 'cream') may be characterized by making the preTerminal Element **extrametrical** (i.e. not counting for stress placement).

Thus, at least by this analysis, the fact that Spanish stress is limited to the final three syllables is an epiphenomenon of no particular theoretical or language particular significance.

In Pirahã, however, there are no lexically assigned stresses. Further, although, as we see directly, Pirahã does have extrametrical syllables these are a very restricted class and can not be appealed to account for any but an infinitesimal portion of proparoxytones. Thus, a theory of Pirahã stress must be able to drive a three syllable domain as the unmarked case.

The immediately following sections attempt to show that neither a grid only account nor a binary branching analysis is able to meet this condition. Following this discussion I argue that a ternary branching analysis can handle the Pirahã facts easily with only minimal modifications of descriptive devices and principles already available in metrical theory.

3 Two failed analyses

3.1 The grid

The grid is unable to recognize the final three syllables as a single constituent, except by stipulation. To see why, consider Prince's (1983, 87) comments on feet:

> "It is a simple matter to divide the grid into feet. A falling or leftheaded foot starts with a $\langle$-level grid entry and runs rightward until another such entry is encountered or the end of a domain. A rising or rightheaded foot is the mirror image opposite."

Since Pirahã stress falls on the last three syllables, foot delimitation must operate from right to left. Now consider the following hypothetical cases. Heavier syllables will register higher on the grid since Pirahã stress is quantity sensitive.[2]

(19) a. CVV. CV. CV [CVV] [CV.CV]
b. VV. GV. CV [VV] [GV.CV]
C. CVV. CV. CV. GV [CVV] [CV.CV.] [GV]

Each of the parsings in (19) fails to recognize that the crucial prosodic domain for Pirahã stress placement is the final three syllables. The grid can only express this as a stipulation, if the statement in (1) is correct. But perhaps the grid can be saved by an alternative conception of Pirahã stress. Let us assume that Pirahã stress is in fact penultimate but that it is subject to the restriction that the penult can only bear the highest level grid mark if neither the ultima nor the antepenult is heavier. If either of these is heavier (thus registering higher on the grid than the penult) the heaviest will be stressed. If both are equal (and heavier than or equal to the penult) the ultima will be stressed. This would be formalizable as in (20)-(22):

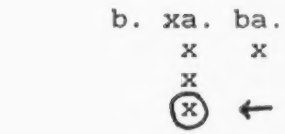
(20) a. baá. ga. hái 'potato' (underlying form)

```
x     x     x
x           x
            x
      (x)
```

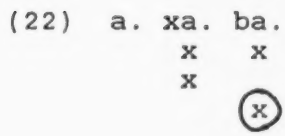
b. baá. ga. hái — Move (x) to ultima

(21) a. xa. ba. gi — 'toucan' (underlying form)

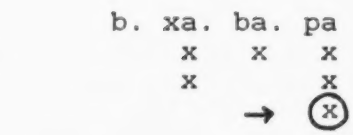
b. xa. ba. gi — Move (x) to antepenult



(22) a. xa. ba. pa — 'proper name' (underlying form)



b. xa. ba. pa — Move (x) to ultima



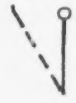
By conceiving of Pirahã stress as primarily penultimate but subject to quantity sensitivity we are indeed able to come up with a description in a grid-like notation. However, such an approach in effect renders the grid meaningless. The layer at which (x) is represented in (20)-(22) bears no principled relation to lower levels of the grid. And yet, at the point where (x) is moved, these lower levels **do** become important, determing the applicability and directionality of the rule (in grid theory, **Move-x**). This global character violates the basic thesis of Prince (1983) that the grid provides a constrained theory of metrical structure by relating each level to immediately preceding levels in a principled manner. Since there is no such relation between levels in (20)-(22), this analysis reduces the grid to the role of a mere descriptive artifice, offering no insights based on independently needed principles and thus explaining nothing. Therefore, it seems that a grid only treatment of Pirahã stress is unavailable.

3.2 Binary trees

Hammond (1984; 1985; 1986) has developed an interesting version of tree theory based on earlier work by, interalia, Halle and Vergnaud (1978; forthcoming) and Hayes (1981). He argues (Hammond 1984, 22ff) that grids alone are inadequate to explain phonological constituent structure (a conclusion forced independently by Pirahã data, as seen above). As already noted, phonological feet are of little theoretical significance in grid theory. But Hammond argues that there are a number of phonological processes which crucially refer to phonological feet. Thus, a theory of prosody must include metrical trees which explicitly recognize the importance of feet. Moreover, in Hammonds tree theory, both rhythmic beat placement (the primary contribution of the grid in earlier metrical analyses, e.g. Liberman and Prince 1977) and phrasing (foot structure) are expressable in a single notation, obviating the need for grids. Hammond's work thus represents an important development and restriction of the formal representations of metrical theory. Finally, Hammond argues that tree structure itself be severely constrained by (i) allowing only binary or unbounded trees and (ii) eliminating all tree structure changing rules except **pruning**.[3]

I argue below that the restriction of trees to binary branching is empirically unjustifiable. Before attempting a tree analysis of Pirahã stress, however, the concept of **obligatory branching** must be introduced. This idea, due originally to Hayes (1981), is interpreted by Hammond (1986) to require, for certain cases in languages in which syllable weight is phonologically relevant, that the terminal, **strong (s)**, nodes may only dominate heavy syllables, whereas nonhead, **weak (w)**, nodes may dominate either heavy or light syllables. In Hammond's model, the **strong** terminals are called **Designated Terminal Elements (DTE)**. To see how obligatory branching works, consider the following Turkish examples (Hammond 1981, 11).

(23) a. Antálya 'city in Turkey'



b. lokánta 'restaurant'

c. Ankara 'city in Turkey'

d. Kenedi 'Kennedy'

In (23a-c), right dominant feet are built from the first heavy (i.e. closed) syllable, in conformity with the obligatory branching parameter. Since (23d) has no heavy syllables, obligatory branching is inapplicable. In these cases, stress defaults to the penult.

(24) Kenedi

Remaining syllables will be attached by the principle of **stray syllable adjunction**:

(25) a. Antálya

b. Iokánta

c. Ankara

d. Kenedi

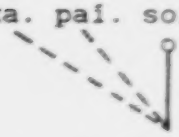
Given these basic notions, let us see how the binary model plus obligatory branching fares in Pirahã. First, as in a grid only account, we must stipulate that the primary focus of stress placement is the penult. To see why, consider how a nonrestrictive algorithm such as (26) would stress the examples in (27) and (28)[4]:

(26) Pirahã stress placement (first restatement):
a. Construct a single obligatory branching, right dominant foot from the right.
b. Join all unattached syllables in a single (unbounded) foot via stray syllable adjunction.

(27) a. paó. hoa. hai 'anaconda/rainbow'

b. xa. ba. gi 'toucan'

c. xa. paí. soí 'shoe'



Once again, though, heavy syllables left of the antepenult cause problems.

(28) a.* poo. gái. hi. aí 'banana'

b.* kao. aí. bo. gi 'jungle spirit'

c.* pia. hao. gi. so. ai. pi 'cooking banana'

To prevent the tree based analysis from stressing the examples in (28) incorrectly, we must stipulate that the penult is the "pivot" for the stress rule.

(29) Pirahã stress placement (second attempt):
- a. Mark the penult as a Designated Terminal Element (DTE)
- b. Shift the DTE to the ultima if it is as heavy or heavier than penult or antepenult.
- c. Shift the DTE to the antepenult if it is heavier than the penult or ultima.
- d. Build a binary tree whose head is the DTE.
- e. Stray syllable adjunction.

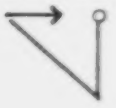
In effect, (29) allows the stress rule to look only to the immediate left or right of the penult, thus limiting stress to the final three syllables. The problem with (29) is that it shifts the DTE without regard to tree structure. This makes it little more than a restatement of the grid treatment in (20)-(22). Consequently, it will face the same problems as the grid. So, let us try again with (30):

(30) Pirahã stress statement (third attempt):
- a. Label the penult as the DTE.
- b. Build a left dominant (binary) tree if the ultima is as heavy or heavier than the penult and antepenult. Otherwise, build a right dominant tree.
- c. Shift the DTE of the resultant tree in accordance with obligatory branching.
- d. Stray syllable adjunction.

As seen in (30c), I am assuming that obligatory branching can be interpreted as a condition on tree representations rather than tree construction. This interpretation enables us to account for the DTE shift by genuine tree geometry principles rather than as a disguised version of Move-x, as in (29). As (31)-(33) illustrate, (30) can indeed describe the facts:

(31) a. paó. hoa. hai (30a) and (30b)

b. paó. hoa. hai (30c)



c. paó. hoa. hai (30d)

(32) a. poo. gáî. hi. aí (30a) and (30b)

b. poo. gái. hi. aí (30c)

c. poo. gái. hi. aí (30d)

(33) a. pia. hao. gi. so. ai. pi (30a) and (30b)

b. pia. hao. gi. so. ai. pi (30c) N/A

c. pia. hoa. gi. so. ai. pi (30d)

Unfortunately, (30) has two serious problems. First, it allows trees in the same domain to branch in either direction. Of course, it is not uncommon to find languages in which foot construction and higher level constituents, such as word trees, are oriented in opposite directions. But to allow trees to branch in either direction at the same level, determined by global[5] properties (e.g. (30d)), would seriously weaken tree theory. Second, this analysis clouds the issue by attempting to relate Pirahã's stress to

apparently similar cases, e.g. Spanish, with penultimate stress. Whereas in these other cases penultimate stress is the most common phenomenon, deviations being explicable via well motivated principles or as lexical exceptions, in Pirahã, penultimate stress is in fact no more common than antepenultimate or ultimate stress. Stipulating the penult as somehow basic thus is an artificial caricature of Pirahã stress. Therefore, it is clear that the restriction of stress placement in Pirahã to the final three sylables, far from being innocuous, presents what appear to be insurmountable theoretical and empirical difficulties for either a grid or a binary tree analysis.

4 Ternarity and obligatory branching

4.1 Amphibrachs

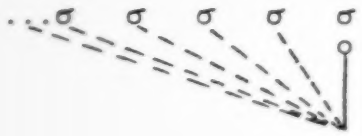
The task of this section is to establish that the problems faced by the grid and binary trees admit of a rather straightforward solution if we assume a particular parametrization of the tree theory conditions outlined by Halle and Vergnaud (forthcoming). The original suggestion for such a parametrization is found in Levin (1985).

Halle and Vergnaud (forthcoming) propose the following.

(34) Conditions on Metrical Structure
a. Every metrical constituent has a head.
b. The head is peripheral.
c. The domain is immediately adjacent to the head.

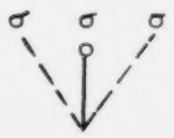
To derive **unbounded** tees (i.e. those with an unrestricted number of branches) Halle and Vergnaud suggest that (34c) be parametrized[6]. However, as Levin (1985, 15) points out, if (34b) were also parametrized, (i.e. if the head need not occur on the right or left) trees with an internal head could also be generated. Levin (1985, 15) claims that this will generate ternary feet of a single type. Unfortunately, this will generate a large number of undesirable structures as well;

(35) Parametrization of (34c) only:

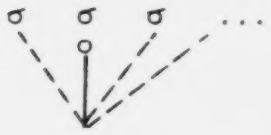
a. 

b.

(36) Parametrization of (34b) only:



(37) Parametrization of (34b) and (34c)

a. 

b.

c.

If Levin's proposal is to be accepted, we must first address the fact that absolutely no languages with head internal, unbounded feet have been documented (and it seems unlikely that any will be).

The simplest solution to our problem would be to disallow either the parametrization of (34c) or of (34b). As a matter of fact, there is no reason to allow (34c) to be parametrized. That is, we need not accept unbounded feet as a primitive of tree theory. Recall that metrical theory must have a rule of stray syllable adjunction. If we require all feet to be binary, except for these produced by the parametrization of (34b), as in (36) above, then "unbounded feet" can be trivially derived by the subsequent application of stray syllable adjunction. This is in fact

what I will assume here (cf. Prince (1985,472) for a similar conclusion).

The structure in (36), produced by the parametrization of (34b) is a particular type of ternary tree, the amphibrach. If we propose that every Pirahã word builds an amphibrach at the right margin (vacuously for disyllables and monosyllables), the three syllable domain of stress placement is automatically accounted for. However, very few words are stressed on the penult, as we have seen. For antepenult and ultima stress the structure in (36) can not be right. Thus, to adequately describe these words, something else is needed. I submit that this "something else" is the obligatory branching parameter, once again interpreting this as a postlexical condition on representations. When more than one syllable in a tree may satisfy obligatory branching, the default value is rightmost. Pirahã stress placement may be stated as in (38).

(38) Pirahã stress placement (final version):
  a. Build an amphibrach at the right margin of the word.
  b. Obligatory branching (default value = rightmost).
  c. Stray syllable adjunction

4.2 Stress in the lexicon

I want to further argue that (38) applies in the lexicon. The evidence for this claim is the fact that secondary stresses arise only in syntactically formed compounds but not in lexical compounds:

(39) xabagi +so**ai**pi 'saw'
toucan - beak

(40) piahaogi +so**ai**pi 'cooking banana'
dolphin - beak

(41) kapiiga +x**ii** +to**ii** 'pencil'
paper - wood - handle

(42) sabáasi +siga**pai**+sai 'syringe'[8]
spirit - piercer - nominalizer

Lexical compounds are recognizable for the facts that (i) the resultant meaning is often much different from the meaning of the compound morphemes, (ii) compounds have the same distribution as single nouns (they may be heads of NPs), (iii) compounds have a single, primary stress. Now consider the examples in (43)-(44) (___ 'primary stress'; 'secondary stress'):

(43) poogáihiái xogabagaí -- poogáihiogabagaí
'want a banana'

(44) kahaí hoáobá --- kahihoáobá
arrow give 'give an arrow'

More examples of the phenomenon illlustrated in (43) and (44), **morphemic fusion**, are discussed in Sect. 5 below. The important observations here are that, unlike lexical compounds, the examples in (43) and (44) function as complete phrases. Their meaning corresponds one-to-one to the meaning of their component words. Finally, **fused morphemes** allow secondary stresses (which words do and do not allow secondary stresses turn out to be a very strong support for ternarity. See Sect. 5 below.).

The simplest way of accounting for the differences just noted is to claim that stress placement applies in the lexicon compounding (thus accenting the result as a single word) but prior to (syntactic) morphemic fusion, where both components of a phrase are accented. It is argued below that rules of resyllabification may alter tree structure in such a way as to shift stress postlexically. But in the absence of such overriding considrations stress will reflect as closely as possible the presyntactic stress pattern. Before considering a wider range of stress facts, it would be worthwhile wo sum up the discussion to this point.

4.3 Summary

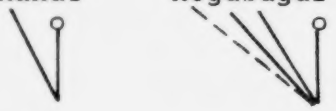
First, it was demonstrated that neither a grid only nor a strictly binary tree analysis is able to express the limitation of Pirahã stress to the final three syllables, without seriously weakening metrical theory. Next, it was shown that by allowing parametrization of the head peripherality condition (but **no** parametrization of the head adjacency condition), amphibrachs may be generated. Finally, I argued that by interpreting obligatory branching as a condition on representations, with a rightmost default value, Pirahã stress placement can be characterized straightforwardly.

In what follows, I will examine a number of the facts which offer independent support for a ternary tree account.

5 Morphemic fusion

In normal speech, in noun + adjective and noun + verb sequences the final vowel of the noun deletes. Also, an initial glottal stop on the following word deletes as well[9]. The entire sequence is then resyllabified from right to left and new tree structure is constructed on the result. The DTE of either word will be shifted after restructuring if needed, via obligatory branching. Whenever two DTEs occur in the same tree following restructuring, the lightest one will be deleted by obligatory branching[10]. A phrase level, right dominant binary tree is constructed over the result.

(45) a. kahaí xogabagaí base forms

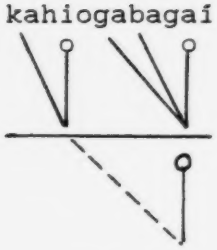


b. kahia metathesis[11]

c. kahi final vowel deletion

d. ogabagaí initial glottal deletion

e. kahiogabagaí restructuring

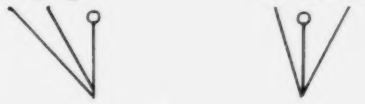


f. phrase tree

'want an arrow'

The derivation in (45) correctly characterizes the fact that in **kahiogabagai** 'want an arrow', the syllable **gai** bears the primary stress and **hio** bears secondary stress. Consider also (46)-(49):

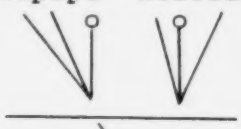
(46) a. xapipái hoáobá base forms



watch give

b. xapipá final vowel deletion

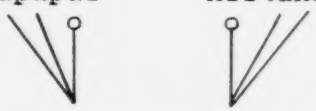
c. xapipá hoáobá restructuring (N/A)



d. phrase tree

'give a watch'

(47) a. xapapaí xiitáha base form

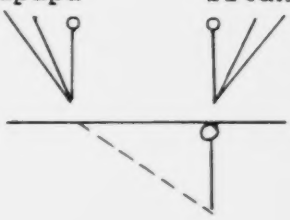


head hurts

b. xapapá final vowel deletion

c. iitáhá initial glottal deletion

d. xapapá iitáhá restructuring (N/A)



e. phrase tree

'head hurts'

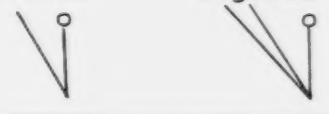
(48) a. baahóísi bagáboí base forms

pig give

b. baahóís final vowel deletion

c. baahóí consonant deletion (see note 9)

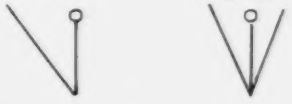
d. baahóí bagáboí restructuring (N/A)



e. phrase tree

'give a pig'

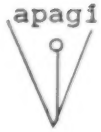
(49) a. tíihíi xapagí[12] base forms



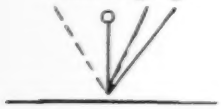
Brazil nut many

b. tiihi — final vowel deletion

c. apagi — initial glottal deletion



d. tiihiapagi — restructuring and obligatory branching



e. — phrase tree (vacuous)

'many Brazil nuts'

In (49), the DTE does not shift off of the final syllable in **tiihii** 'Brazil nut' because resyllabification has the effect of replacing the deleted /i/ with the initial /a/ of **xapagi** 'much/many'. Tree construction from right to left will then place 2 DTEs within a single tree, **pa** and **hia**. Obligatory branching will require that only **hia**, the heaviest of these two potential DTEs, be interpreted as the DTE (cf. note 13). Since the **DTE** is transmorphemic, involving also the same morpheme as the first syllable of the phrase, **tii**, this syllable will simply be added to the whole by stray syllable adjunction, rather than bear a secondary stress.

Morphemic fusion of nouns and disyllabic or monosyllabic adjectives and verbs is interesting here since it allows us to test the ternarity hypothesis further. This is so since if ternary trees are rebuilt from right to left, the DTE of an adjective or verb will frequently occur in the same tree as the DTE of the preceding noun. As (50)-(54) show, obligatory branching plus ternary trees makes the correct predictions.

(50) a. sitoí hoí — base forms

egg two

b. sito — final vowel deletion

c. sitohói — restructuring

d. sitohói — oblilgatory branching[13]

e. phrase tree

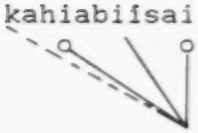
'two eggs'

(51) a. kahiaí biísai — base forms

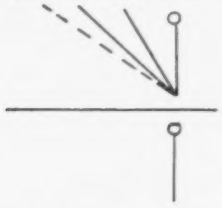
basket red

b. kahia — final vowel deletion

c. kahiabiísai — restructuring



d. kahiabiisai — obligatory branching



e. phrase tree

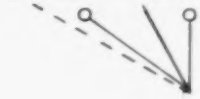
'red basket'

(52) a. poogáihiai hói — base forms

banana one

b. poogáihia — final vowel deletion

c. poogáihiahói — restructuring



d. poogáihiahói — obligatory branching

e. phrase tree

'one banana'

(53) a. xisigíhii hiaba — base forms

meat negative

b. xisigíhí — final vowel deletion

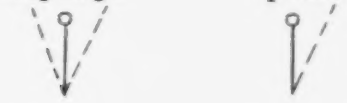
c. xísigíhihiaba — restructuring

d. xísigíhihiaba — obligatory branching

e. — phrase tree

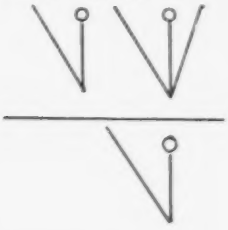
'have no meat'

(54) a. kapiiga piixi — base forms



paper long

b. kapiigapiixi — restructuring/obligatory branching



c. — phrase tree

'long paper'

In (54) two levels of stress are heard even though the following adjective **piixi** 'long' is disyllabic. This is because the DTE of **kapiiga** 'paper' is to the left of the first ternary tree formed from the right following restructuring.

To summarize: in noun + adjective and noun + verb sequences, the DTE of either the noun or following word will delete if it occurs in the same tree as a heavier DTE. Only

a theory allowing ternary trees is able to predict when two DTEs will appear in the same tree. Further, this phenomenon of DTE deletion can not be analyzed as a purely linear restriction against adjacent DTEs since, as examples like (47) show, DTEs may in fact be adjacent if they are in separate trees. This evidence serves to support the proposal in Sect. 4 above that Pirahã stress placement operates within ternary trees.

Now let us conclude this study with a brief look at the interaction of extrametrical syllable and ternary trees.

6 Extrametricality

The following words appear to violate the stress rule given in (38) above.

(55) xoiboi**bii**sai 'species of fish'

(56) ko**hoi**biisai 'proper name'

(57) siga**pai**sai 'syringe'

(58) kapiigakaga**kai**sai 'calculator'

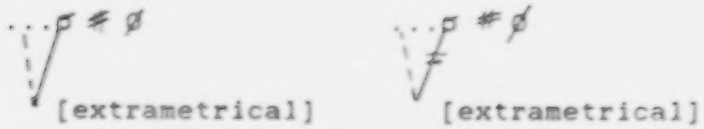
(59) xiixi**boi**tisai 'pencil sharpener'

According to our stress rule, each of these words should be stressed on the final syllable. Interestingly, this is the same in all the examples, the nominalizer, **-sai**. The obvious suggestion is that **-sai** is extrametrical (Hayes 1982). However, **-sai** is still relevant for determining the domain of stress placement, since if it were not, example (55) would be stressed as in (60).

(60) * **xoi**biobiisai

To account for this, we can simply analyse extrametricality as a form of tree pruning (**prior** to obligatory branching).

(61) Extrametricality



By (61), (55) can be derived as in (62):

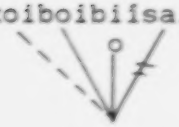
(62) a. xoiboibiisai nominalized form

[extrametrical]

b. xoiboibiisai tree structure

(amphibrach)

c. xoiboibiisai pruning



Extrametricality (-**sai** is the only such syllable I am aware of) thus is consistent with the hypothesis that the domain of Pirahã stress is determined by lexically constructed ternary trees in that extrametrical syllables, although themselves unstressable, still are relevant to determining the final three syllable domain of rule (38).

The extrametrical rule claims that an extrametrical syllable will only be pruned when phrase final. To verify this, consider the interaction of morphemic fusion and extrametricality in (63):

(63) a. xitisai xogabai

bitter medicine wanting

b. xitisa final vowel deletion

c. ogabaí initial glottal deletion

d. xitísaogabaí restructuring

'wanting bitter medicine'

(63) shows that an extrametrical syllable can be stressed in a nonphrase final, restructured environment.

7 Conclusion

In this paper it has been argued that the analysis of Pirahã stress placement makes three important contributions to phonological theory:

(i) Onsets may be relevant for stress placement,
(ii) Ternary trees are a necessary enrichment of metrical theory and are derivable from the parametrization of Halle and Vergnaud's (forthcoming) **head peripherality condition**,
(iii) Obligatory branching (as revised by Hammond (1986)) may be interpreted as a condition on representations rather than tree construction per se.

Notes

* I want to thank the Pirahã, especially **Kohoibiíhai** and **Xaópisi**, for putting up with me for these past eight years. My wife Keren has been of inestimable help in discussing, criticizing, and correcting. Juliette Levin, Michael Hammond, Morris Halle, Jay Keyser, Alan Prince, and Moira Yip have all encouraged me to think through the issues here more carefully. High tone is represented as /'/, low tone as /O/. Glottal, /?/, is written as **x**.

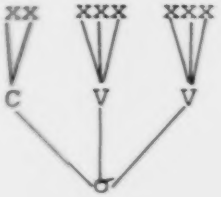
1. This of course implies that notation is significant as a characterization of the properties of **Universal Grammar**, a traditional assumption of generative linguists.

2. In (19) I am assuming that each syllable weight will be considered for foot demarcation, since **heaviest** in (19) is determined relative to neighboring syllables in a given word. Alternatively, we could attempt to scan the word from the right, adding a -level grid mark (**x**), to the first

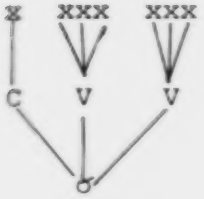
occurence of the heaviest syllable type in the word. This would correctly parse (19a) and (19b) but in examples of the (19c) type (cf. (15)-(18) above) the result would be incorrectly given as [CVV.CV.CV.GV].

3. That is, all trees must be restricted to just two branches or unrestricted.

4. For obligatory branching to work in Pirahã, it must be revised so as to pick out the heaviest syllable type from the five valued scale in (2) above. This could be done in a variety of ways but the most likely seems to be to determine syllable weight via internal branching complexity as in (i) (where long consonants are mapped to two positions on the timing tier - cf. Levin (1985) and Clements and Keyser (1983)). Here I use Levin's (1985) **X** notation:

(i) a. CVV = 

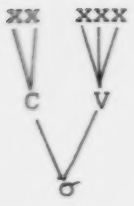
(11 branches)

b. GVV = 

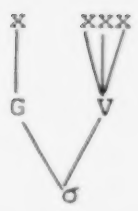
(10 branches)

c. VV =

(8 branches)

d. CV = 

(7 branches)

e. GV = 

(6 branches)

5. Hammond (1985, 422ff) discusses a case of bidirectional foot construction in Lenakel verbs:

(i) a. Make the final syllable extrametrical.
b. Build one right-dominant binary foot on the right margin.
c. Build left-dominant binary feet left to right.

Notice though, that this bidirectionality is restricted to opposite margins of the word and is constant for all Lenakel verbs. This is much different than (30). In (30) it is necessary to look at the final three syllables **before** tree construction and determine tree directionality according to surface level accent, a **global constraint**, depriving directionality of any explanatory force.

6. This part of my discussion is essentially a paraphrase of Levin (1985, 14).

7. Note that a word tree will not be necessary.

8. The final syllable in (42), **-sai** is extrametrical and thus not stressable. See Sect. 4.7.

9. When final vowel deletion results in an unsyllabifiable CC sequence, the leftmost C is deleted as in:

(i) xagi báaxi ⟶ xagobáaxái ⟶ xabáaxái
path good 'good path'

the rules are ordered as in (ii)

(ii) a. /V/ ⟶ O /___ $\neq$C
b. C ⟶ O /___C

10. We cannot simply derive **new** DTEs in conjunction with resyllabification or restructuring since this would fail to capture the fact that only postlexically fused morphemes but **not** lexical compounds bear secondary stresses. In other words, DTEs are assigned lexically but are subject to the **postlexical** obligatory branching condition. This is one

more support for the conception of obligatory branching as a condition on representations.

11. All **sai** and **hai** syllables undergo metathesis before final vowel deflection. See Sheldon (1974) for discussion.

12. The DTE of **xapagi** will be deleted presumably because of resyllabification. It can not shift to the penult **pa** because obligatory branching has not yet applied. Thus DTEs are affected only by resyllabification and obligatory branching.

13. In (50d) the high tone of the previous DTE syllable shifts to the final syllable. Cf. Sheldon (1974) for one analysis of this phenomenon. High tone in effect "follows" the stress in certain morphemic combinations. I have no analysis of this high tone shift here. Only "perturbable" adjectives and verbs undergo this. In Hammonds (1984) theory of DTEs, only one is allowed per foot. Obligatory branching will require that the DTE appear on the heaviest syllable. Any other DTE which appeared with the foot prior to reconstruction will be redefined by obligatory branching as a nonDTE, weak node. Again, in the event that two (or three) potential DTEs are of equal weight, the default value will select the rightmost.

References

Clements, G.N. and S.J. Keyser. CV Phonology. Cambridge, MA: MIT Press.

Davis, S. 1985. Syllable weight in some Australian languages. In: M. Niepokuj, et.al. eds.

Everett, D. 1985. Syllable weight, sloppy phonemes, and channels in Pirahã discourse. In: M. Niepokuj, et.al. eds.

_____ 1986. Pirahã In: D. Derbyshire and G. Pullum eds. Handbook of Amazonian languages. Berlin: Mouton.

_____ and K. Everett. 1984. On the relevance of syllable onstes to stress placement. LI. 15. 705-711.

Halle, M. and J.R. Verguand (forthcoming). An essay on English stress.

Hammond, M. 1984. Constraining metrical theory: a modular theory of rhythm and destressing. Bloomington: Indiana University Linguistics Club.

_____ 1985. Main stress and parallel metric planes. In: M. Niepoluj, et.al. eds..

_____ 1986. The obligatory branching parameter in Metrical theory. NLLT. 4.

Harris, J. 1983. Syllable structure and stress in Spanish: a nonlinear analysis. Cambridge, MA: MIT Press.

Hayes, B. 1981. A metrical theory of stress rules. Bloomington: Indiana University Linguistics Club.

_____ 1982. Extrametricality and English stress. LI. 13. 227-76.

_____ 1985. Iambic and trochaic rhythm in stress rules. In: M. Niepokuj, et.al. eds.

Hyman, K, 1985. A theory of phonological weight. Dordrecht: Foris.

Levin, J. 1985a. A metrical theory of syllabicity. MIT doctoral dissertation.

_____ 1985b. Generating ternary feet. Unpublished MS. University of Texas, Austin.

Liberman, M. and A. Prince. 1977. On stress and linguistic rhythm. LI. 11. 511-562.

Niepokuj, M. et.al eds. 1985. Proceedings of the eleventh annual meeting Berkeley Linguistics Society.

Prince, A. 1983. Relating to the grid. LI. 14. 19-100.

_____ 1985. Improving tree theory. In: M. Niepokuj. et.al. eds.

Sheldon, S. 1974. Some morphophonemic and tone perturbation rules in Mura-Pirahã. IJAL. 40 279-282.

REDUPLICATION IN MAJANG

Pete Unseth

1 Introduction
2 Reduplication
  2.1 Formation of reduplicated verb stems
  2.2 Uses of reduplicated verbs
    2.2.1 Intensive action
    2.2.2 Repeated action
    2.2.3 Reciprocal action
    2.2.4 Present continuous
    2.2.5 Imperfect future
    2.2.6 Third person plural
    2.2.7 Derived forms
  2.3 Other reduplicated words
  2.4 Pseudo-reduplicated forms
  2.5 Comparative data
3 Prefix tV-
  3.1 Grammatical uses of tV-
  3.2 Derived nouns based on prefixed verbs
  3.3 Semantic aspects of tV-
  3.4 Restrictions on the affixation of tV-
  3.5 Potential ambiguities with stem initial /t/ and reduplication
  3.6 Comparative data
4 Summary

1 Introduction

Majang is a member of the Nilo-Saharan family of languages, part of the Surma cluster. It is spoken in southwestern Ethiopia in Kefa and Illubabor provinces, with a small group also in Wellega.[1] There are only two published linguistic descriptions, one by Cerulli (1948) and a recent, more comprehensive and reliable one by Bender (1983).

In his article, Bender gives five examples of verb reduplication in Majang (1983a:121). Building on Bender, this present paper gives many more examples of verb reduplication, explains the different uses of reduplication, and describes the phonological rules for reduplicating. Examples of (possibly) reduplicated forms from other parts of speech are also given. Comparative data is also given to show that some reduplication was also found in Proto-Surma. In section 3, a verb prefix tV- is described, which may be confused with reduplication, since it covers some of the same semantic range and can appear in forms phonologically similar to reduplication.

2 Reduplication

As mentioned above, Bender briefly discussed the

function of reduplication in Majang verbs. Many other words also appear to be reduplicated. The phonological rules for verb reduplication (and nouns derived from verbs) are clear, but other words that appear to be reduplicated can reduplicate by different rules. The term "reduplication" is sometimes used in this paper in a broad sense to include any word that repeats a syllable or more, even if there is no clear evidence of a non-reduplicated root.

Sapir noted that "The process [of reduplication] is generally employed, with self-evident symbolism, to indicate such concepts as distribution, plurality, repetition, customary activity, increase of size, added intensity, continuance," (1921:76). In a cross-language study of reduplication, Moravcsik observed that "The most outstanding single concept that reduplicative constructions recurently express in various languages is the concept of increased quantity ... quantity of referents and amount of emphasis." (1978:317). Both of these observations are in harmony with Bender's observation that reduplicated verb forms in Majang are used to denote "intensification and related notions." (1983a:121). Reduplication in Majang, however, goes beyond this to include some tense and aspect paradigms, as well.

2.1 Formation of reduplicated verb stems

Before giving any examples of reduplicated verbs, a brief listing of verb root types is given below. (Since reduplication follows the same pattern for long and short vowels, this listing of syllable structures does not distinguish long and short vowels.) Verb roots can be one or two syllables, in a variety of syllable shapes, the most common being CVC, such as **kɛt-** 'chop' and **baal-** 'dance'. Other patterns include CV **gii-** 'grind grain', CVCVC **jumur-** 'answer', CVCCV **turku-** 'invert', CVCV **tiro-** 'sneeze', CVCC **ʔorp-** 'give a name', CVCVC **moomon-** 'weed', and CVCCVC **kurkudʕ-** 'tie'.

The reduplication pattern in all examples of verb reduplication is predictable, (with a possible exception given below). The initial consonant-vowel sequence of the final syllable of the stem is reduplicated.[2] The same rule also explains the formation of Bender's duplicated verb forms. The reduplicated form carries the same tone as the stem. (To avoid breaking up syllables in the morpheme-by-morpheme glosses, reduplication will be treated as being prefixed to the final syllable, but this does not mean that this matter is settled.) The following examples demonstrate this reduplication process with monosyllabic and disyllabic roots.

If the stem is monosyllabic, the initial consonant-vowel sequence of the stem is reduplicated as a prefix, as in **jojokiikan** 'they wounded each other' from **jok-**. If the stem is disyllabic, the initial

consonant-vowel sequence of the final syllable is reduplicated as an infix, as in **bˀokokotiikan** 'they killed each other" from **bˀokot**.

So far, I have found only one CVCC stem, **ʔorp** 'give a name'. For such stems, assuming there are others, this rule will have to be modified, since it reduplicates as **ʔororopaaŋ**. In this case, the penultimate consonant reduplicates with a vowel that is not found in the stem, though it is an echo of the stem vowel.[3]

(1) **Reduplicated monosyllabic roots:**

gloss	**root**	**3s present continuous**
wash	ʔaad-	ʔaaʔaadŋ
chop	kɛt-	kɛkɛtŋ
grind grain	gii-	giigiiŋ

(2) **Reduplicated disyllabic roots:**

weed	moomon-	moomomoŋ
invert	turku-	turkukuŋ
tie	kukudˀ-	kukudˀɪŋ
answer	jumur-	jumumurŋ

2.2 Uses of reduplicated verbs

Verbs marked by reduplication have a number of distinctive uses. These are listed below. Some of these are obviously related, such as intensity and repetition and reciprocity. This paper attempts to list etic categories, rather than emic. Some of the examples given are from Bender (1983), but these have all been checked.

2.2.1 Intensive action. Reduplication can be used to mark action as being more intense: **ŋuuŋuulikoŋ** 'he broke it to bits', as opposed to simple **ŋuulikoŋ** 'he broke it', from the stem **ŋuul**.

2.2.2 Repeated action. Repetition or habitual action can be marked by reduplication, with the meaning "do all the time" or "do again and again".

(3)

gloss	**root**	**reduplicated**	**translation**
eat	dˀaam-	dˀadˀami	'eat all the time'
speak	ton-	totoni	'speak all the time'
bend	yik-	yiyikiŋ	'sway back and forth'
bring	diil-	didilkɔɔŋ	'bring all the time'[4]

Reduplication to mark repeated action can also be used in the imperative: **kɛtɪk** 'chop!' (from **kɛt-**) is reduplicated to **kɛkɛtɪk** 'chop repeatedly!'. Also, **yawok** 'slice!' (from **yaw-**) is reduplicated to **yayawok** 'slice repeatedly!'.

This reduplication is also used in negative imperative forms:

(4) kati yawɪn! 'don't slice'
kati yayawɪn! 'don't slice repeatedly'

Reduplication is also found with structures similar to infinitives, e.g. **dˀeaa yayaw** 'I want to slice repeatedly.' Such reduplication is also used in finite forms, such as the past tense, **yayawkɔŋ** 's/he sliced repeatedly'.

Reduplication for intensive and repetitive can also be found in negative indicative sentences. This is significant since many verbs have irregular forms when they are negated (see Unseth (forthcoming).)

(5) Ku-ii-ko ya-yaw-e.
NEG-1p-PST DUP-slice-NEG
'We did not slice repeatedly.'

(6) Ku-ii-ko ŋuu-ŋuul-e.
NEG-1pl-PST DUP-'break'-NEG
'We did not shatter.'

2.2.3 Reciprocal action. Reduplication is used to mark verbs for reciprocity, the idea that "they did it to each other". The morphology is marked on the following example. (with **n3** meaning "non-third person" and **IV** meaning "intransitive"):

(7) ko-koñ-ik-ar-ko-ŋ
DUP-help-n3-2p-PST-IV
'You (pl) helped each other.'

The following examples are all 3rd person plural, past tense, meaning 'they did it to each other':

(8)	**gloss**	**root**	**reduplicated**
	killed	bˀokot-	bˀokokotiikoŋ
	wounded	jok-	jojokiikoŋ
	gave to	gij-	gigijiikoŋ
	loved	por-	poporiikoŋ
	hated	bˀoy-	bˀobˀoyiikoŋ
	helped	koñ-	kokoñiikoŋ
	called	riy-	ririyikoŋ
	fought	tim-	titimiikoŋ
	saw	dɛn-	dɛdɛniikoŋ
	knew	dˀigir-	dˀigigiriikoŋ
	insulted	ñon-	ñoñoniikoŋ
	washed	ʔaad-	ʔaaʔaadiikoŋ

Reciprocal action can also be marked on 1st and 2nd person plural, as well

(9) 1st per: **totonikiikoŋ** 'we conversed with each other'
kokoñikiikoŋ 'we helped each other'
2nd per: **totonikarkoŋ** 'you conversed with each other'
kokoñikarkoŋ 'you helped each other'

Reduplication is also used to mark imperative reciprocals. The root for 'love' is **por-**, which reduplicates to form **poporikoŋ** 'they loved each other'. The imperative 'love each other' is the reduplicated form **impoporike**. (The simple 2nd person plural imperative 'love!' is **impore**.)

Reduplication to mark recipocal action is also found in the negative, as in the following:

(10) Ku-ɛr-ko wo<-no->noy-it.
NEG-3p-PST exchange-DUP-NEG
'They did not exchange with each other.'

2.2.4 Present continuous. Reduplication is also used to mark a present continuous action when the action is in progress as it is said, such as "I am eating now." This reduplication can occur on both transitive and intransitive verbs, in all persons and numbers. Examples are given below:

(11)

gloss	**root**	**reduplicated**	**translation**
go	ŋaar-	ŋaaŋaarŋ	'He is now going.'
crack	laal-	laalaaliŋ	'It is now cracking.'
thatch	ʔutum-	ʔututumiŋ	'They are now thatching'
kill	bʕokot-	bʕokokorŋ	'He is now killing.'[5]
beat	bol-	bobolaaŋ	'I am now beating.'
split	palat-	palalataaŋ	'I am now splitting.'
chop	kɛt-	kɛkɛtaaŋ	'I am now chopping.'
err	tik-	titikaaŋ	'I am now erring.'

Note that the last three examples may be confused with the **tV-** prefix discussed in section 3.5. In this case, the initial **ti-** is a duplication of the stem, not the **tV-** found elsewhere.

Reduplication for present continuous is also found in negated verb stems.

(12) gii-gii-k-aa-ŋ
DUP-grind-n3-1s-IV
'I am now grinding.'

K-aa gii-gii-t.
NEG-1s DUP-grind-NEG
'I am not grinding.'

Reduplication to show present continuous action is also found in the passive voice. Compare the active and passive sentences below:

(13) gii-gii-O-ŋ
DUP-grind-3s-IV
'S/he is now grinding.'

gii-gi-O-e-ŋ
DUP-grind-3s-PSV-IV
'It is now being ground.'

2.2.5 Imperfect future. The grammatical label of this category is admittedly tentative. These forms indicate that the action has not yet been done, but that it will be done in the future, such as 'I have not yet eaten,' but implying "I will eat." This is not a true negative; in a true negative construction, the negative verb is **ku** and true negated verbs are postposed, often in a suppletive form. (see Unseth (forthcoming) for description of negatives.)

(14)

Gloss	Root	Reduplicated	Translation
converse	ton-	intotoniŋ	'He hasn't yet conversed.'
argue	palam-	inpalalamiŋ	'He hasn't yet argued.'
thatch	ʔutum-	inʔututumiŋ	'He hasn't yet thatched.'
grind	gii-	ingiigiiyiŋ	'He hasn't yet ground.'
eat	dˤam-	indˤadˤamiŋ	'He hasn't yet eaten.'

As further evidence that this is not a negative, no verb form is postposed to the end of the sentence, as in the following example:

(15) Inkɛkɛt kɛɛt.
3s-chop tree
'He hasn't yet chopped a tree.'

2.2.6 Third person plural. Cerulli claimed to find reduplication in the verb morphology marking third person plural subject, but gave no examples to illustrate this (1948:141,142). The only third person plural examples found in his data show no sign of this, only a sufix **-ar-**, which is exactly what Bender later described (1983a:131). Cerulli may have misunderstood **-k-**, a directional marker, in his examples, such as **melarkiko** 'they came' (1948:142).

Bender found the same grammatical form, but demonstrated that the directional **-k-** and person marker **-ar-** have metathesized, a process found is several verbs of motion (1983:131,132). The underlying morphemes of Cerulli's "**melarkiko**" are as follows:

(16) mɛl-k-ar-ko
come-DIR-3p-PST

This example of third person plural, then, is not, after all, reduplication. The first **-k-** is a directional marker, the second part of the past tense morpheme.

Cerulli may also have been misled by a sentence like:

(17) Ku-k-u-ko mɛlɛ.
NEG-DIR-3s-PST come
'He did not come.'

Cerulli's discovery of reduplication in the third person plural is therefore to be disregarded.

2.2.7 Derived forms. The Majang form for 'blood' is a loan, **yɛrum**.' From this noun, Majang has derived a verb **yɛrum**, 'bleed'. This verb, derived from the noun, reduplicates by the usual patterns: **yɛrurumŋ** 's/he is bleeding'.

Derived nouns can also be built from reduplicated verb forms:

(18) kɛkɛtkan 'chopper' (one who chops a lot of wood)
yayawkan 'slicer' (one who slices a lot)

2.3 Other reduplicated words

Many other words appear to be reduplicated, including nouns, adjectives, some verb roots, and adverbs.

Several other nouns which have repetitive syllables have not been traced to non-repetitive forms. These include: **diidi** 'scar', **kooko** 'snake' **(also found in Suri and Didinga)**, **wowoj** 'frog', **jiiji** 'ancestor', **loolomi** 'charcoal'.[7] 'finger' **polpol** (compare 'fingernail' **copolkoi** and Me'en 'finger' **kobolkit**).

Reduplicated kinship terms are common throughout the world, such as "mama" and "papa". In Majang, these include: 'mother (of second person)' **ŋaaŋa** , 'aunt' **maame**, 'Daddy' **baabey**, 'father (of 2nd person)' **bˤaabˤa** (the differrence in the articulation of the bilabial stops on the two terms for 'father' was confirmed in careful checking).

One reduplicated adverb has also been found, **agurgur** fast/quickly', as in **reeri agurgur** 'He ran fast'.

A few verbs have synchronic roots that may represent diachronic duplication, though unduplicated roots have not been found. Cerulli gave the verb 'to tie' as **kurkur-**, apparently a fully reduplicated verb stem (1948:161), though my Majang contact rejects this, giving **kurkudˤ-**, which Cerulli had listed as an alternate).

The verbs 'enter' **guuguñ-**, 'weed' **moomon-**, and 'rub' **moomoj-** all appear to be reduplicated, but cannot be reduced to a non-duplicated form. On the contrary, they are subject to the usual reduplicating patterns: **guuguguñŋ** 's/he is entering' and **moomomonŋ** 's/he is weeding' and **moomoomojaaŋ** 'I am rubbing'. The noun for 'weed' is related to this apparently reduplicated verb root, **moomonan**.[8]

Some reduplicated intransitive adjectival verbs have been found. Two of them are translated (into Amharic) as 'soft': **ñulñul** (for surfaces), **kokoŋ** (for foods). Two other reduplicated adjectives were found, **dˤidˤik** 'right (hand)' and **dɔdɔŋ** 'wet' (the -ŋ being a suffix). Other verbs in this category include **kalkal** 'be dangerous' and **dˤodˤoy-** 'exceed in height'.

2.4 Pseudo-reduplicated forms

Not surprisingly, various sequences of similar sounding morphemes can combine to produce sequences of similar syllables that look like reduplicated words.

(19) ku-O-ku-ko tar diilɛ.
NEG-3s-DIR-PST meat bring
'He did not bring meat.'

kaaka-katok-e tak
cave-PL-LOC in
'in caves'

2.5 Comparative data

Within Surma, explicit descriptions of reduplication have been found in two other languages, Me'en and Murle. In Me'en, Ricci found verb reduplication used to express "repetition or intensity of an action," (1972:184). He found reduplication of the beginnings of stems, the ends, or entire stems. (The limited number of examples and the seeming lack of systematic change in meaning on the reduplicated stems raises doubts about how productive and widespread the process is in Me'en.) This reduplication does not follow so regular a pattern nor is it so integrated into the grammar as in Majang.

In Murle, Arensen discovered "the intransitive marker in the imperfect is a reduplication of the final consonant of the verb stem," (1982:85). He also uses the term "reduplication" to describe certain other morphemes in Murle, but the term is not appropriate. In his subjunctive and perfect examples, it is better analyzed as vowel harmony. He speaks of reduplication of consonants in the imperfect, but it is consonant assimilation (1982:63-69). There does appear to be at least one reduplicated verb root in Murle, **liŋliŋ** 'work'. The Murle noun **riiriton** 'reflection' is a probable reduplicated form. Cognate non-reduplicated forms are documented in three other Surma languages: Majang **rii** 'shadow (of animate)', Didinga **lori** 'shadow, and Mursi **rii** 'shade'.

There is no description of reduplication in other Surma languages, though, not surprisingly, other Surma languages have words that appear to be reduplicated. Didinga, for instance, has 'idly' **takaktak**, 'divination by pebbles'

tooto, 'cry of alarm' **lulu**, **mamai** 'maternal uncle', etc., (Driberg 1931). The Tirma form for 'cold' is **walala** and Suri 'claw' is **tɛrtɛra**. There is not enough data available to say anything beyond the observation that such seemingly reduplicated words exist.

Though Turton and Bender did not specifically discuss reduplication in Mursi, there is at least one example of reduplication in the data. The singular form of the adjective 'big' is **bwi**, which reduplicates to **bibi** when modifying a plural noun (1976:555). This reflects the tendency of Surma languages to mark adjectives for plural.

This form preserves a Proto-Surma pattern of reduplication for marking number on the adjective 'big'. This gives at least one solid example that a reduplication process was found in Proto-Surma. The plural of the adjective 'big' also shows (diachronic) reduplication in Majang (Northern Surma) and Me'en and Mursi (Southern Surma):

	singular	**plural**
Majang	obii	bobɛr
Me'en	buitiui	biibii-da
Mursi	bwi	bibi
Murle	appi	
Shuro	buitini	
Suri	abbo	
Tirma	abbu	

Singular and plural of 'big' in Surma languages

The above chart suggests a Proto-Surma form of the approximate shape ***abi** or ***abu** meaning 'big', with a reduplicated plural form.

3 Prefix tV-

Majang has a prefix **tV-** which sometimes covers areas of meaning similar to reduplication. Also, on verbs with initial **tV-** sequences, it mimics reduplication. The vowel of the prefix is left unspecified since it always harmonizes totally, becoming identical in point of articulation to the first vowel of the verb stem. Examples of this prefix are presented here, with the verb in the imperative form, with and without the prefix. The imperative suffix is **-Vk**, the vowel usually harmonising (at least partially) to the stem. This, together with the harmony of the prefixing process, produces some three syllable words having the same vowel in each syllable.

gloss	prefixed imperative	root
dig	tokowuk	kow-
fear	tubˤuruk	bˤur-
help	tokoñok	koñ-
cry	tukudˤuk	kudˤ-
cough	takanɪk	kan-
touch	tukumuk	kum-
fly/jump	tipirik	pir-
pierce	tumukuk	muk-
bite	takawuk	kaw-
chop	tɛkɛtɪk	kɛt-
swim	teleyik	ley-
flee	togonok	gon-
call	tiriyik	riy-
return	tumurak	mur-

Forms with tV- prefix

3.1 Grammatical uses of tV- prefix

The tV- prefix appears in a limited number of grammatical environments, the most common being the imperative. However, in negative imperatives, the prefix is lost:

(21) Tɛ-kɛt-ɪk! 'chop a lot!'
Kati kɛt-ɪn! 'don't chop!'
*Kati tɛ-kɛt-ɪn! A prefixed form is not acceptable.

I have not been able to elicit any prefixed past tense forms. Some verbs are prefixed to form infinitives. It is interesting that the verbs with prefixed infinitives are often the same verbs with prefixed negated stems. These prefixed infinitives include: 'to chop' **tɛkɛ** (from **kɛt-**, verbs often lose final consonants in the infinitive), 'to call' **tiriy**, 'to laugh' **togo** (from **rog-**), 'to flee' **togon**, and 'to swim' **tiliiy**.

The prefix can also appear in the negated form of some verbs, (though some negated verb stems are suppletive). For example, the stem for 'call' is **riy-**, as in **riyŋ** 'he calls'. But in the negative, the prefix **ti-** is obligatory:

(22) Ku-ko ti-riy.
NEG-PST tV-call
'He did not call.'

It is not permissible to say ***kuko riy**. Other verbs with **tV-** obligatorily prefixed to the negated stem include 'swim' **liiy-**, negative **tiliiy**, 'cry' **kudˤ-**, negative **tukudˤ**, and 'flee' **gon-**, negative **togon**.

Verbs cannot be both reduplicated and prefixed with tV- at the same time. For example, the root for 'fear' is **bˤur**,

often prefixed with tV-, as in **tubˀuruk** 'fear!'. However, 'He is now fearing' is **bˀubˀurŋ**, reduplicated, but cannot be prefixed. A prefixed reduplicated form, such as ***tubˀubˀurŋ** is not acceptable.

3.2 Derived nouns based on prefixed verbs

Some nouns are derived from the prefixed form of the verb. The verb 'help' has the stem **koñ-**, which is sometimes prefixed with tV-, in this case **to-**. The derived noun 'help, assistance' is built on the prefixed verb, **tokoñ**. The derived agent 'helper' is also built on the prefixed form, **tokoñtan**.

The verb stem **ɯɛy** 'roast (grain)' usually appears with the prefix **tɛ-**. The product form is based on the prefixed form of the verb: **tɛɯɛytanak** 'roasted grain', (**-tan** marks product, **-ak** plural). Similarly, the agent 'sorcerer' is **tamaltan**, from the verb root **mal-** 'do sorcery'.

The word 'name' is **tiriya**, based on the prefixed form of 'call' **riy**. The noun 'gift' is **tigi**, from the prefixed form of the verb **gij-** 'give'. The verb root 'fear' is **bˀur**, often prefixed with tV-. The noun 'fear' is a prefixed form, **tubˀur**, as in **lakaa tubˀur** 'I have fear' ('I am afraid').

3.3 Semantic aspects of tV-

The function of this tV- prefix is far from clear. When translating into Amharic, a Majang language helper suffixed an object marker onto the Amharic verb form for several tV- forms, such as 'fear', 'dig', 'bite', 'swim'. The prefixed form for 'pull out' was given with an Amharic causative prefix. These forms suggest some sort of transitivizing function.

Deliberately asking Majang speakers to clarify distinctions between prefixed and non-prefixed forms elicited a variety of responses. One person suggested that the prefixed **tokowuk** would mean 'dig it out!', as oppossed to the non-prefixed form meaning simply 'dig!'. Another man said that the prefixed form **tipirik** meant 'jump far!'. The prefixed form **tukumuk** 'pierce!' was given with the example of throwing a spear, piercing an object at some distance. One person said that he would use the non-prefixed form **koñok** 'help!' with a family member, rather than use the prefixed **tokoñok**. The imperative of 'give' was given as **gijik** (stem **gij**), but 'give me!' was given as prefixed **tigit**, (the **-t-** is a first person singular object marker (Bender 1983:128), but the loss of the stem-final consonant is unexplained). The imperative for "ferment (mead)!" is **yaakɪk**, but the prefixed **tayakɪk** means "ferment it later!" (the shortening of the vowel in the stem of the prefixed form is unexplained). There is no obvious thread of meaning

connecting all of these explanations. They suggest that the prefix indicates some sort of intensity, but this is far from clear. The prefix is no longer productive in Majang, but it is still preserved in many words (like the English prefix **con-**).

3.4 Restrictions on the affixation of tV-

This prefix can only be affixed to a restricted set of verbs. In a random check of 50 verbs, only eleven could be prefixed. Those that can be prefixed include both transitive and intransitive verbs. Forms of the verb 'to be' have not been found with **tV-**. This restriction on the verbs that can be prefixed is also discussed in sec. 3.6.

3.5 Potential ambiguities with stem initial /t/ and reduplication

The presence of the **tV-** prefix is not always obvious. Firstly, some verb roots begin with **t-** followed by a vowel. Secondly, some reduplicated forms resemble prefixed forms.

In the first category, many verbs have initial **t-** followed by a vowel, but in these, the initial sequence of **t-** and a vowel is part of the root, not removeable. These **t-** initial verbs include, 'sneeze' **tiro-**, 'err, make a mistake' **tik-**, 'begin' **tuuk-**, 'hear' **tey-**, 'pierce' **tim-**, 'roast (meat)' **tuuj-**, 'open (a door)' **tay-**, 'mediate' **taak-**, and 'be thin' **tɛŋɛɛl-**.

The fact that the initial **t-** is part of the stem, not the prefix, is especially clear if the vowel of the first syllable is different than the vowel of the second, as in **tiroik** 'sneeze!'. This is also clear if the first syllable ends in a consonant, such as **turkuk** 'invert'.

The second possible ambiguity is for monosyllabic verbs with a root of the shape **tV(C)** to reduplicate in such a way as to resemble a **tV-** prefixed form. This looks the same as the **tV-** prefixed forms of the same verbs, since the vowel is always the same as the stem vowel in both cases.

(23)

gloss	root	prefixed	translation
pierce	tim-	titimaaŋ	'I am now piercing.'
hear	tiy-	titiyarɪŋ	'They are now hearing.'
err, lose	tik-	titikiŋ	'He is now losing, erring.'

For these forms, the meaning indicates it is a case of reduplication. However, for the imperatives 'pierce!' **titimik** and 'err, lose repeatedly!' **titikik**, it is not obvious if these are examples of reduplication or **tV-** prefixation, since both are allowed for imperatives, such as **yayawɪk** 'slice repeatedly!' (clearly reduplicated from **yaw-**) and **tɛkɛtɪk** 'chop repeatedly!' (clearly a **tV-** prefixed form from **kɛt-**).

It is possible to determine if a form is reduplicated or prefixed with **tV-** by forming a negative imperative. The **tV-** prefix is not allowed in the negative imperative (see Sect. 2.1.1), but it is allowed for reduplicated forms, such as **kati kɛkɛtɪn** 'don't chop repeatedly!'. This test shows that **titimik** 'pierce repeatedly!' is a not a prefixed form but a reduplicated form. Since the negative imperative can be **kati titimin** 'don't pierce repeatedly!', this must be reduplicated, since only reduplicated forms (but not **tV-** prefixed forms) are allowed in negative imperatives.

3.6 Comparative data

There is a description of a **tV-** verb prefix in at least one other Surma language. It appears in Didinga, where it is also difficult to define precisely. Odden analyzed it as part of a "completive stem" (1983:162) It is not (yet) possible to determine if these are genuine cognates, but the resemblance in phonological shape and in function is striking. Odden wrote, "My basic hypothesis regarding **tV-** is that it is a lexically governed tense-aspect prefix: (note that the **tV-** prefix appears only in the completive). There were a few cases where I would like to treat **tV-** as simply the first syllable of a polysyllabic stem ... [It generally seems to] correlate with something like ... specific, definite, complete." (personal communication 1985). In Didinga, as in Majang, the vowel of the prefix is identical to the following stem vowel.

One difference between Didinga and Majang in the use of tV- is that **tV-** is not found on negated verbs in Didinga (Odden 1983:162), but it is in Majang (see 2.1).

In both Didinga and Majang, the prefix is found only on a restricted set of verbs. Of the small set of verbs that take the **tV-** prefix in Odden's examples, at least one seems cognate with a prefixed form in Majang. Odden gives the "completive stem" **tEŋEd** for 'chop', from **ŋEd-**, (1983:162). In Majang, the root for 'chop' is **kɛt-**, prefixing to form **tɛkɛt-**. Bender's unpublished Suri notes give **tɛŋɛra** for 'cut (wood), which may also represent a **tV-** prefix on a cognate stem.

There is also evidence of a **tV-** prefix in Me'en. Ricci lists **ta-** as a probable prefix in the form **tabadaboy** 'to smear' since he found probable cognates without the **ta-** (1972:396,397). He also gave 'to blow' with two forms, based on the root **bai-**, **tabaidai** and **babaiboi** (1972:449). This also suggests a prefix **tV-**.

Though it is not possible to exactly equate the **tV-**prefixes in the these languages, Odden's suggestion of "complete" or "definite" is intriguing and not incompatible with the Majang examples.

3 Summary

In summary, this paper has demonstrated how Majang reduplicates words for a variety of functions, not only for intensification, but also for tense and reciprocal action.

This discussion of Majang reduplication has also demonstrated that reduplication was used in Proto-Surma, and has given some directions for the study of reduplication in Proto-Surma and other Surma languages.

This paper has also presented the first evidence on the Majang verb prefix tV-, differentiating it from reduplication. Evidence has also been presented that this morpheme was found in Proto-Surma.

Notes

1. The data for this paper was gathered in Ethiopia between August 1984 and March 1986, while I worked under the Institute of Ethiopian Studies, Addis Ababa University. In addition to the many who have cooperated at AAU, I wish to acknowledge the cooperation of a number of local officials who have made field work possible. On one field trip, Anbessa Tefera was sponsored by the Institute of Language Studies Research and Publications Committee. Special thanks to the Majang people who have provided the data and answered my many strange questions. This paper has benefitted from comments by E.-A. Gutt, R.J. Sim, H.-G. Will, J. Clifton, and C. Kutsch Lojenga. M. L. Bender kindly provided unpublished material on Surma languages. My supportive wife, Carole, makes it all possible.

Bender listed six vowel phonemes for Majang (1983). I have now become convinced of a vowel distinction based on Advanced Tongue Root, as Odden has found in Didinga (1983). However, since my reanalysis of the vowels and my sorting out of grammatical and lexical tone is not yet complete, these two important phonological points are regretfully omitted from this paper.

The abbreviations should be clear from the examples. The digraphs bˀ and dˀ represent voiced implosive stops.

For descriptions of relevant syntactic matters, such as negative sentences and their sometimes suppletive stems, see my forthcoming syntax paper.

2. Since the reduplicated CV is inserted next to the original CV, it is not clear whether the reduplication is inserted before the final syllable or whether it is inserted just after the original CV sequence (just before the stem final consonant, if any).

3. The above gives a summary of the general rules for

reduplication in Majang. If more stems with final CC clusters are found (in addition to the single example ʔorp- 'give a name'), Majang will be an interesting language for the application of CV theory, such as Marantz's work (1982). Otherwise, the rules of reduplication for Majang are in harmony with the general principles of reduplication in other langauges (Moravcsik 1978).

4. I cannot explain why the long vowel of the stem becomes shortened in reduplicated form. In checking, I confirmed that it is genuinely short in the reduplicated form, but genuinely long in the stem form.

5. The final consonant of 'kill' becomes [r] by a morphophonemic rule that changes stops to liquids before a nasal. The underlying stem final /t/ of 'kill' is found when there is a following vowel, as in **bʕokokotaaŋ** 'I am now killing'.

6. Fleming cites the origin of this form in the Omotic Majoid languages (1983:544).

7. The Majang form for 'charcoal' may have originated not by reduplication, but by assimilation. The Murle form for 'charcoal' is **zolomo-**. In Majang, the initial /z/ could have been assimilated to the /l/ within the stem, resulting in **loolomi**. However, the opposite process, deriving the Murle initial /z/ by dissimilation is also possible.

8. This may be cognate with the form **mana** 'field' in both Murle and Didinga, as well as some Eastern Nilotic languages (Dimmendaal 1982:104). Since weeds quickly take over a field in the Majang swidden type of agriculture (Stauder 1971:26ff), a semantic shift from 'field' to 'weed' is very plausible.

References

Arensen, Jon. 1982. Murle grammar. (Occasional Papers in the study of Sudanese Languages, 2.) Juba: College of Education, Summer Institute of Linguistics, and Institute of Regional Languages.

Bender, M. Lionel. 1983a. Majang phonology and morphology. Bender, ed., 114-147.

--------ed., 1983b. Nilo-Saharan Languages Studies. (Monograph 10, Committee on Northeast African Studies.) East Lansing: African Studies Center, Michigan State University.

Cerulli, Enrico. 1948. Il Linguiaggio dei Masongo nell'Etiopia Occidentale. Rassegna di Studi Etiopici, 7.2:131-166.

Dimmendaal, Gerrit. 1982. Contacts Between Eastern Nilotic and Surma Groups. Culture History in the Southern Sudan, J. Mack and P. Robertshaw, eds., 101-110. (Memoir 8.) Nairobi: British Institute of Eastern Africa.

Driberg, J.H. 1931. The Didinga Language. Mitteilungen des Seminars fuer Orientalishce Sprachen, 34:139-182.

Fleming, Harold. 1983. Surma Etymologies. Rainer Vossen and Marianne Bechhaus-Gerst, eds., Nilotic Studies, 523-555. Proceedings of the International Symposium on the Languages and History of the Nilotic Peoples, Cologne, 1982. (Koelner Beitraege zur Afrikanistic, Band 10.) Berlin: Dietrich Reimer Verlag.

Marantz, Alec. 1982. Re Reduplication. Linguistic Inquiry, 13.3:435-482.

Moravcsik, Edith. 1978. Reduplication Constructions. Word Structure, 297-334. (Universals of Human Language, ed. by Joseph Greenberg, vol 3.) Stanford, CA: Standford University Press.

Odden, David. 1983. Aspects of Didinga Phonology and Morphology. Bender, ed., 148-176.

Ricci, Lanfranco. 1972. Materiali per la Lingua Mekan. Rassegna di Studi Etiopici, 25:90-455.

Sapir, Edmund. 1921. Language. New York: Harcourt, Brace, and World.

Stauder, Jack. 1971. The Majangir. (Cambridge Studies in Social Anthropology, 5.) Cambridge: Cambridge University Press.

Turton, David and M.L. Bender. 1976. Mursi. The Non-Semitic Languages of Ethiopia, M.L. Bender, ed. 533-561. (Committee on Ethiopian Studies, Monograph 5.) East Lansing: African Studies Center, Michigan State University.

Unseth, Pete. forthcoming. A Sketch of Majang Syntax. In Bender, ed. Nilo-Saharan Language Studies, 2 (tentative title).

Indirect Objects and Incorporation in Mazatec[1]

Daniel Agee and Stephen Marlett

1 Introduction
2 Overview of Mazatec
3 Indirect objects and incorporation
4 Benefactive Advancement
5 2-3 Retreat
  5.1 General 2-3 Retreat
  5.2 Governed 2-3 Retreat
6 Inversion
7 3-chomeurs
8 Conclusion

1 Introduction

This paper deals with indirect objects in Mazatec. It has two fairly modest goals. The first is to describe the superficial characteristics of indirect objects in Mazatec. We claim that they occur as prepositional phrases which are obligatorily incorporated into the verb (in a specific way described below). The second goal is to describe the constructions in which indirect objects occur in Mazatec. We do this primarily within the framework of Relational Grammar (Perlmutter 1983a, Perlmutter and Rosen 1984) claiming that indirect objects occur in simple clauses and also in clauses with Benefactive-3 Advancement, 2-3 Retreat (personal and impersonal), and (impersonal) Inversion. 3-chomeurs, which sometimes occur, are marked differently from final indirect objects. All of these claims are novel for Mazatec. In addition, this is the first time for which such constructions have been posited for any Otomanguean language.

2 Overview of Mazatec

The Mazatec spoken in San Jeronimo Tecoatl (district of Teotitlán de Flores Magón), Oaxaca, Mexico, is similar in many ways to the other Mazatec languages of the area. Nevertheless, the facts and analyses discussed in this paper do not necessarily apply to the other languages. We first present some typological characteristics.

Clausal word order does not appear to be as fixed in San Jeronimo Mazatec (henceforth SJM) as in many

other Otomanguean languages (which are most commonly VSO). It is prepositional, adjectives (quantifiers excluded) follow the head noun, and the possessed noun precedes the possessor.

The preposition 'of' (which figures importantly in the rest of this paper) precedes most possessors, although most body parts and some kinship terms do not require the preposition 'of' before the possessor. (As we see later, it also introduces indirect objects.) This preposition is an enclitic phonologically and has two shapes, depending on the person of its complement. It is **la̲** when the complement is second person singular or third person, and **na̲** otherwise, as shown in (1). A simple example is given in (2).

		shape	complement
(1)	'of'	la̲	2s, 3
		na̲	1s, 1pi, 1pe, 2p

(2) a. koyo̲ hmá **la̲** ʻilióʻ
horse black of
'Julio's black horse'

b. koyo̲ hmá **na̲ -hẽ̲**
horse black of 1pe
'our black horse'

SJM is similar to the Zapotec languages in that it has two sets of pronouns--clitic and nonclitic, which are described below-- but (unlike Zapotec) it also has person agreement with the subject. (We do not discuss here the issue of whether it is the initial or final level that is relevant to this agreement; we do not have the necessary range of data yet to discuss this important issue.) This person agreement is shown primarily by stem changes, and is of a very limited type. The most usual pattern is for there to be two stems, one of which is used when the subject is first person singular or third person, and the other stem used otherwise. The stem change may be fairly minor, as in (3a), or radical, as in (3b).

		1s, 3 stem	2s, 1pi, 1pe, 2p stem
(3) a.	'see'	kocehẽ	ʻicehẽ
	'cry'	khindaya	ʻhindaya
	'throw'	sikathẽ	nikathẽ
	'work'	siʻšá	niʻšá

b.	'talk'	ĕha	nokhoa
	'give'	cha	?vi
	'take'	?va	ĕ?a

A few verbs have more stems and show the agreement more fully.

(4) a.	'go'	hvi	1s, 3 stem
		?mi	2s stem
		onki	1pi, 1pe, 2p stem
b.	'say'	šă	1s stem
		co	3 stem
		si	2s stem
		višó	1pi, 1pe, 2p stem

Stem glosses indicate subject agreement in the examples below, but not all of the ambiguity is expressed in the translations.

The nonclitic pronouns are given in (5) and the clitic pronouns in (6). We distinguish these pronouns in these charts and in the examples below by glossing them differently, as shown.

(5) Nonclitic pronouns

?ă	I, me
hi	you
he	she, he, him, her
hña	we, us (inclusive)
hĭ	we, us (exclusive)
hño	you (plural)
hña	they, them

(6) Clitic pronouns[2]

-a	1s
-i	2s
-ø	3 (singular or plural)
-á	1pi
-hē	1pe
-o	2p

The nonclitic pronouns occur in various positions. They are used to give emphasis to the subject, as in (7), or to the direct object, as in (8).

(7) Mahē, **?ă** ko-niya -a
no I FUT-wash=dishes 1s
'No, I'll wash the dishes.'

(8) s-kot?a -a **he**
FUT-test/1s,3 1s him
'I will test him'

Nonclitic pronouns are also required for the object of most prepositions along with the clitic pronoun. In (9) both are necessary. (10) illustrates that the preposition **la/na** 'of' requires only the clitic pronoun.

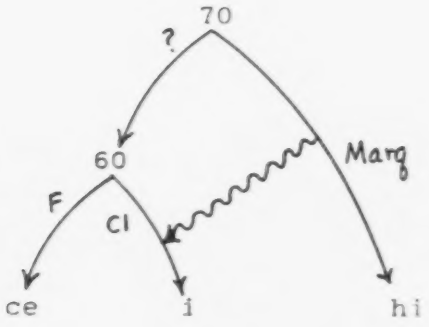
(9) ñánda **ce -i hi**
where of 2s you
'Where are you from?'

* ñánda ce -i

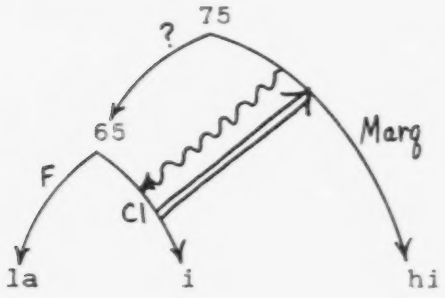
(10) koyo **la -i**
horse of 2s
'your horse'

We propose that prepositional phrases such as the one in (9) have the structure shown in (11a) and that prepositional phrases such as the one in (10) have the structures shown in (11b). We discuss the structure in Arc Pair Grammar terms (Johnson and Postal 1980). In both (11a) and (11b) the Marquee arc sponsors a clitic arc which shares a tail with the Flag arc. The sponsor relation is indicated in these diagrams by a wiggly arrow. The relation which node 60 bears to node 70 is left undetermined. (In phrase structure terms, node 60 would most likely be labeled P'.) In (11b), however, the clitic arc erases the Marquee arc. The erase relation is indicated in this diagram by a double arrow. The conditions for this erasure are only when a pronoun heads the Marquee arc. Therefore, in the surface structure of prepositional phrases with 'of', as in (11b), only one nominal appears overtly following the preposition: either the clitic pronoun (first or second person), or the non-pronominal complement. This complex prepositional phrase structure is relevant to later discussion.

(11) a.



b.



The clitic pronouns are used in the first NP position following the head of verb phrases (as subject), as shown in (12), of noun phrases (as possessor), as shown in (13), and in prepositional phrases, as seen above in (9-10).

(12) ñánda onki **-ō**
where go/1p,2p 2p
'Where are you (plural) going?'

(13) nithð **-a**
nose 1s
'my nose'

3 Indirect objects and incorporation

SJM is similar to English and other languages in that surface 3s in SJM sponsor prepositional phrase structures. The preposition **la/na** (which was glossed as 'of' above, but which is not glossed below) is the flag in such phrases in SJM, just as **to** is the flag in English. SJM is different from English, and many other languages, in two ways. First, the structure of prepositional phrases in

SJM is different, as we have seen above. Second, the preposition-clitic structure of surface 3s is obligatorily incorporated into the verb. Examples (14a-b) are clauses with surface 3s, and example (14c) is with a surface Comitative; note that the prepositions **la** and **ko** do not occur contiguous to their NP complements. The preposition-clitic structure is phonologically dependent on and not separable from the verb. This is represented by a hyphen before both the preposition as well as the clitic.

(14) a. he khi -**la** -**ø** karta **čilič**
she write/1s,3 3 letter
'she writes a letter to Julio'

b. he sikathĕ -**la** -**ø** pelota **čihveli**
she throw/1s,3 3 ball
'she throws the ball to Felix'

c. ti-síšá -**ko** -**ø** -a **Alberto**
CNT-work/1s,3 with 3 1s
'I am working with Alberto'

The examples in (15) show that not only the preposition is incorporated, but the preposition-clitic complex. This fact is the motivation for the constituent structure shown in (11).

(15) a. he khi -**la** -**i** karta
she write/1s,3 2s letter
'she writes a letter to you'

b. kó co -**la** -**i** Alma
how say/3 2s
'what did Alma say to you?'

c. kó co -**na** -**a** Alma
how say/3 1s
'what did Alma say to me?'

In the following sections we describe other constructions in which surface 3s occur in SJM.

4 Benefactive advancement

Benefactives optionally advance to 3 in SJM. The examples in (16) have final Benefactives. The examples in (17) are of Benefactives which are final 3s. As final and surface 3s, these nominals in (17) are flagged with **la**/**na**.

(16) a. Alma ti-vinta komida **nka -a ʔã**
CNT-make food for 1s me
'Alma is making food for me'

b. kohve ti-vithe̲ čiliõ **nka -ø čihveli**
coffee CNT-plant for 3
'Julio is planting coffee for Felix'

(17) a. Alma ti-vinta **-la̲ -ø** komida **čiliõ**
CNT-make 3 food
'Alma is making food for Julio'

b. čiliõ kohve ti-vithe̲ **la̲ -ø Ramõ**
coffee CNT-plant 3
'Julio is planting coffee for Ramon'

c. čisi ti-sĩšá **-na̲ -a**
CNT-work/1s,3 1s
'Cecilia is working for me'

5 2-3 Retreat

5.1 General 2-3 Retreat

We identify many clauses in SJM as involving 2-3 Retreat. First, when the 2 of a transitive stratum is not third person, it obligatorily retreats to 3. Such a retreat might be merely one way of enforcing a language-specific constraint of the following type: If a̲ heads a final 2-arc, then a̲ is third person. We do not give the ungrammatical sentences here to fully document that such a constraint exists. Since no grammatical sentence with a first or second person as final 2 exists, there is no pattern on the basis of which one can proceed to construct examples to test.

Since the initial 2 is a final 3 in the examples in (18),[3] the preposition **la̲** occurs in (18) but not in (19).

(18) a. s-kini **-la̲ -i**
FUT-send/1s,3 2s
'I will send you'

b. s-kot?a **-la̲ -i**
FUT-test/1s,3 2s
'I will test you'

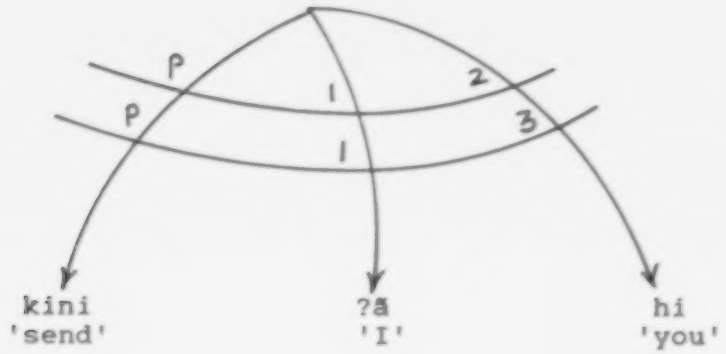
c. s-kot?a **-na̲ -a**
FUT-test/1s,3 1s
'he will test me'

(19) a. s-kot?a -a Ramŏ
FUT-test/1s,3 1s
'I will test Ramon'

b. s-kot?a -a (he)
FUT-test/1s,c 1s him
'I will test him'

We claim that the structure of (18a) includes the following:

(20)



5.2 Governed 2-3 Retreat

There are a few verbs, both transitive and intransitive, which virtually require 2-3 Retreat, regardless of whether the direct object is third person or not.[4] The transitive verbs that pertain to this group include **hvithinki** 'follow', **konačă** 'trick', and **?nčé** 'hear'. Examples are given in (21). As (22) shows, if there is no 2, there is no 2-3 Retreat and therefore no preposition **la**.

(21) a. čihveli ti-hvithinki -**la** -**ø** **Maria**
CNT-follow 3
'Felix is following Mary'

b. čiliho kis-konačă -**la** -**ø**
DP-trick 3
'Alfonso tricked him'

c. miki ?nčé -**la** -**ø**
not hear/1s,3 3
'I don't hear him/her/it/them'

(22) miki ?nc̣é -a
not hear/1s,3 1s
'I don't hear'

The group of intransitive predicates which require 2-3 Retreat all determine initially unaccusative strata and include **nihña** 'tired', **hva** 'awaken', and **iho** 'hungry'. Examples are given in (23).

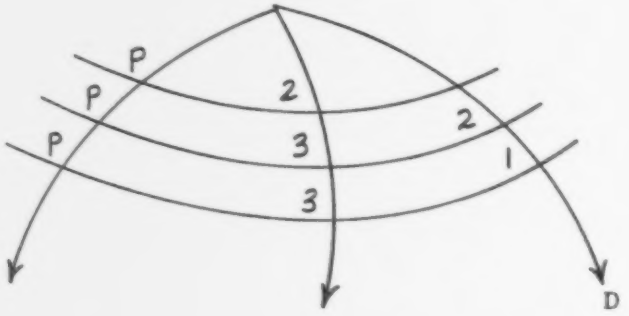
(23) a. hie nihña -**la** -**i**
already tired 2s
'you are already tired'

b. kó sí ko-hva -**la** -**i**
how do/1s,3 RP-awaken 2s
'how did you awaken?'

c. a ihó -**la** -**Ø čihveli**
Q hungry 3
'Is Felix hungry?'

We propose that these clauses involve a silent dummy, as shown in (24).

(24)



This analysis requires some explanation. This dummy enters as a 2; it overruns the initial 2, inducing the demotion of the 2, but not putting it en chomage. The overrun relation is proposed in Johnson and Postal (1980). Perlmutter (1986) states it as follows:

> If two arcs A and B have the same tail and the same R-sign, and B's first coordinate index is +1 of A's last coordinate index, then B overruns A.

This relation is used for various purposes. Most importantly here, it is used by Berinstein (1986) in a reformulation of the Active Dummy Law. The Active Dummy Law, referred to in Perlmutter (1983b), and discussed in

Perlmutter (1983c), is modified by Berinstein to state as follows:

> A dummy must overrun some nominal if its departure stratum contains a nuclear term arc.

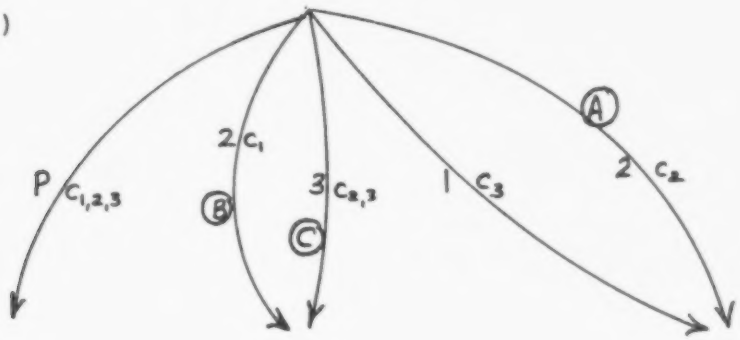
Our analysis is consistent with Berinstein's formulation of this law. The stratum preceding the entrance of the dummy contains a nuclear term arc, and the dummy overruns it.

The analysis shown in (24) is also consistent with the Final 1 Law and the Entrant Term Law. The latter is a new proposal by Bickford (1986); his tentative formulation is given below:

> For each pair of arcs A and B with tail c, where A is an entrant term arc [i.e., its head does not head a distinct term arc at a previous stratum in the clause in question] and B is a term arc, if the last coordinate of B is less than the first coordinate of A, then there is an entrant term arc (either A or another arc) which overruns B.

The diagram below is a relational network equivalent to the stratal diagram in (24).

(25)



According to the definition given above, arcs A and B are entrant term arcs, but arc C is not. The last coordinate of arc B (1) is less than the first coordinate of arc A (2). The Entrant Term Law claims that arc A (or some other arc) must overrun B, as it does.

6 Inversion

A small class of verbs in SJM_5 idiosyncratically occur primarily in Inversion clauses. The initial subject

occurs only as a final indirect object. Predicates which belong to this class include the verbs **mehē** 'want' and **sasē** 'like'. The examples in (26) illustrate these verbs with nominal direct objects, and those in (27) illustrate them with clausal direct objects.

(26) a. mehē -**na** -**a** hnko čho
want 1s one radio
'I want a radio'

b. **Alma** sasē -**la** -**ø** nohmá
like 3 beans
'Alma likes beans'

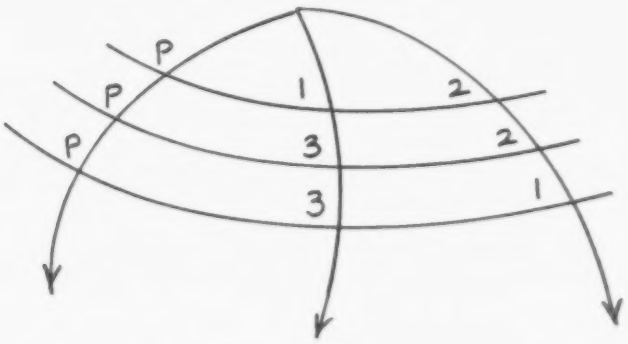
(27) a. sasē -**la** -**ø** khē **čihveli**
like 3 eat/1s,3
'Felix likes to eat'

b. mehē (-**la** -**ø**) khē **čihveli**
want 3 eat/1s,3
'Felix wants to eat'

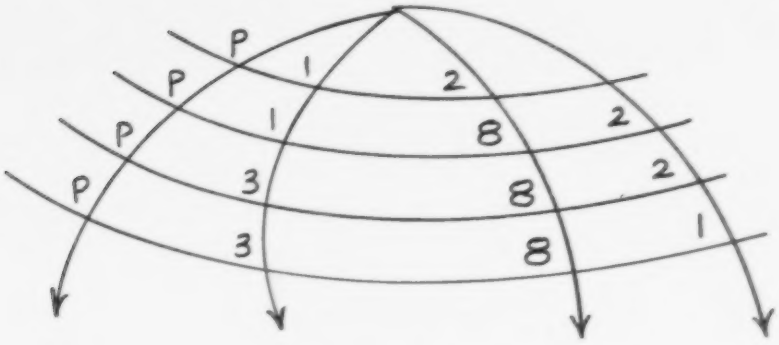
The optionality of **la** in (27b) may indicate that Inversion is possible but not necessary under certain conditions.

It is unclear at this time whether the examples in (27) are of personal inversion or impersonal inversion. Both analyses are represented in (28). The impersonal inversion analysis given in (28b) is again consistent with the Active Dummy Law, the Entrant Term Law, and the Final 1 Law. In section 6 we discuss examples in which Inversion interacts with 2-3 Retreat; in those cases the clause is most certainly impersonal.

(28) a. Personal Inversion



b. Impersonal Inversion



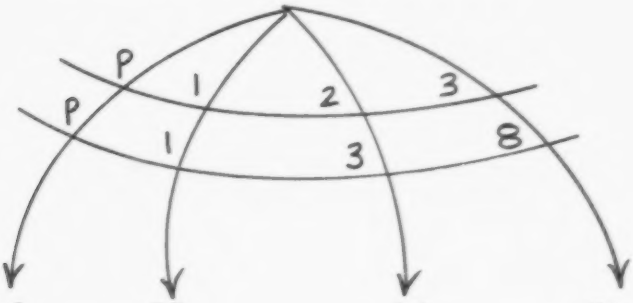
7 3-chomeurs

It is not difficult to imagine that there exist in SJM sentences in which more than one indirect object must occur, given the structures posited above. We discuss two types of these sentences in this section.

An example with an initial indirect object and also first or second person initial 2 is given in (29). The structure which (29) is claimed to have is given in (30).

(29) s-kini **-ko** **-∅** -la -i **čisi**
FUT-send/1s,3 3 2s
'I will send you to Cecilia'

(30)



The retreat of the initial 2 to 3 puts the initial 3 en chomage. The initial 2 is the final 3 and incorporates as expected. The preposition associated with the 3-chomeur is **ko** 'with' rather than **la**, and is incorporated. The following informal generalization can be made (anticipating some facts which follow):

(31) A nominal b which heads an 8-arc having a

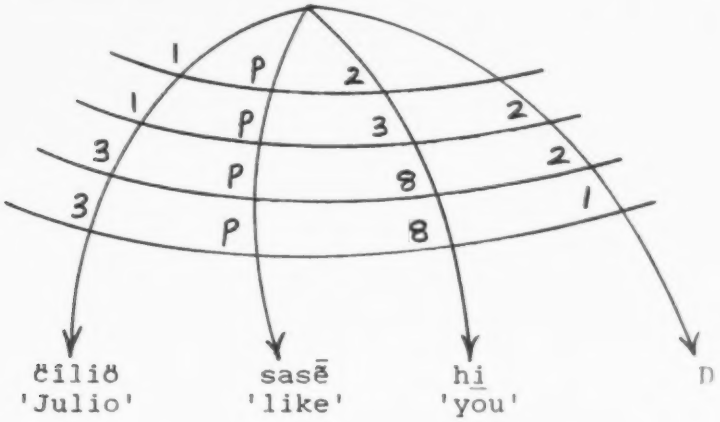
3-arc predecessor in the first stratum is flagged with **ko**.

This flagging rule therefore crucially makes reference to both the final stratum and the initial stratum.

Another situation in which 3-chomeurs are expected to arise is when the initial 2 of an Inversion predicate is not third person. Both Inversion and 2-3 Retreat occur. An example is given in (32) and the proposed analysis in (33).

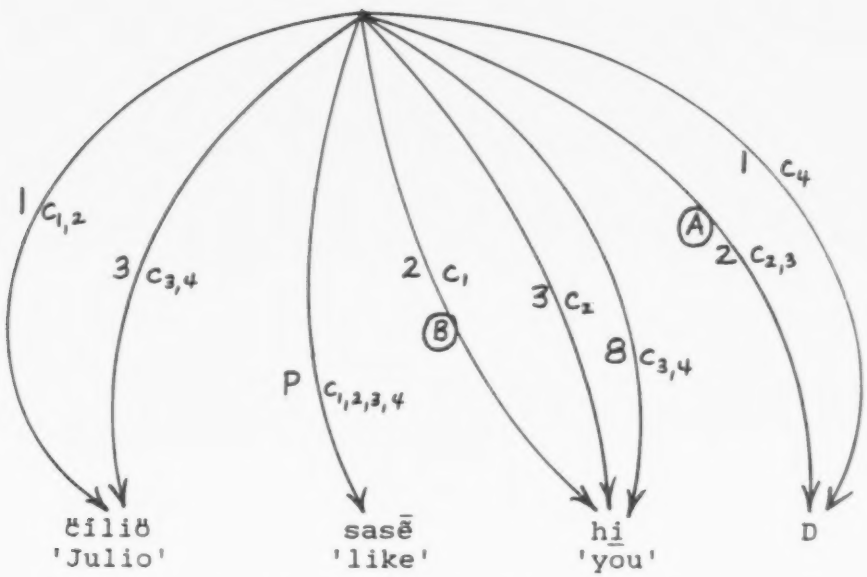
(32) čilið sasē -la -Ø **nka -i hi**
like 3 2s you
'Julio likes you'

(33)



The retreat of the initial 1 to 3 puts the non-initial 3 en chomage. It is not necessary to stipulate that 2-3 Retreat occurs 'before' Inversion in (32) since the 'reverse' order would violate the Entrant Term Law. To see that this is so, consider first the relational network equivalent to (33).

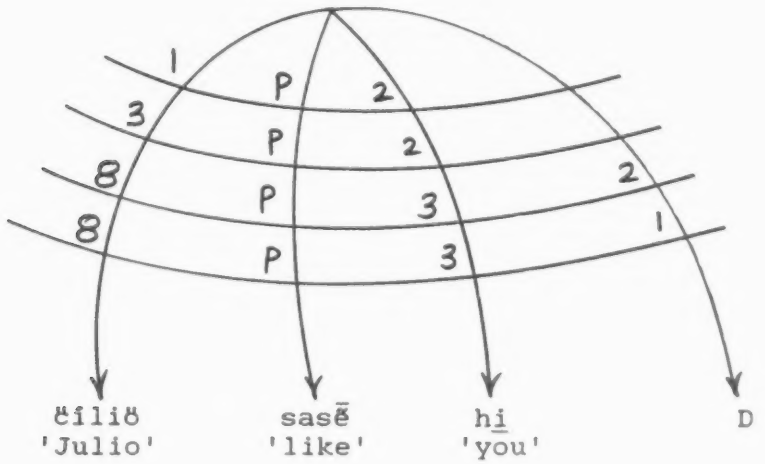
(34)

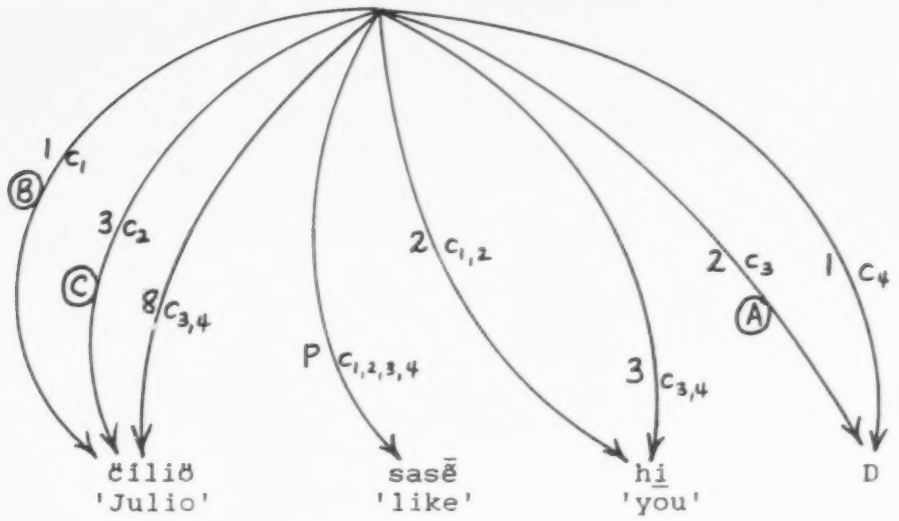


Arc A and Arc B are the relevant arcs, just as in the discussion of (24) above. This structure satisfies the Entrant Term Law in the same way.

Structure (35), in which the 'reverse' order is attempted, violates this law since the dummy does not head an arc which overruns an arc headed by 'Julio'.

(35)





In (35) Arc A is the relevant entrant term arc. Since Arc A does not overrun Arcs B and C, this structure violates the Entrant Term Law. The Mazatec facts are therefore explained by the interaction of these principles.

In (32) the 3-chomeur is flagged with the preposition **nka** 'for', which does not incorporate. The following generalization can be stated:

(36) A nominal b which heads an 8-arc having a 3-arc local predecessor and a remote predecessor which is not a 3-arc is flagged with **nka**.

This flagging rule makes reference to both the final stratum and two nonfinal strata.

8 Conclusion

At first blush many clauses in Mazatec appear to be strange and unlike those of languages with which we are more familiar. It has been shown here that these characteristics result from a combination of fairly familiar constructions and phenomena: (a) prepositional phrases, (b) incorporation, (c) advancements and retreats, and (d) impersonal constructions. The prepositional phrases have two uncommon properties: (a) the internally complex structure, and (b) the suppletive allomorphy of the preposition 'of'. The incorporation facts are interesting because (a) incorporation is obligatory, and (b) the preposition-complex, rather than simply the preposition, is incorporated. Once

we take notice of these common phenomena and some of their less common characteristics, Mazatec begins to look more familiar.

We have claimed that many sentences in Mazatec contain surface 3s which do not correspond to what have come to be called 'notional indirect objects'. The surface 3s in Mazatec correspond to intial 3s, initial Benefactives, initial 2s, and initial 1s. The conditions under which 3s are claimed to occur sometimes interact, in predictable ways, with the result that 3-chomeurs occur. The fact that proposed laws such as the Entrant Term Law and the Active Dummy Law make correct predictions for the Mazatec facts provides support for these laws.

Notes

1. Some topics discussed in this preliminary report were presented at the Conference on Relational Grammar and Grammatical Relations held at Ohio State University, Columbus, Ohio, in May of 1986, and at a technical forum of the Summer Institute of Linguistics in Tucson, Arizona in March of 1986. We thank the audiences at these presentations for their comments and questions. We also thank Peter Constable and Albert Bickford for their helpful suggestions on earlier drafts of this paper.

The following abbreviations are used: 1s -- first person singular, 2s -- second person singular, 3 -- third person, 1pi -- first person plural inclusive, 1pe -- first person plural exclusive, 1p -- first person plural (inclusive or exclusive), 2p -- second person plural; CNT -- Continuative, DP -- Distant Past, FUT -- Future, RP -- Recent Past, (unmarked tense/aspect = timeless aspect), Q -- question word. The following tone markings are used: a̲ -- 4 (low) tone, á -- 1 (high) tone, á̲ -- rising tone (either 4-2 or 3-1), a -- 3 tone, ā -- 2 tone. The following abbreviations are used in relational networks and stratal diagrams: Cl -- Clitic, D -- Dummy, F -- Flag, Marq -- Marquee, P -- Predicate, 1 -- subject, 2 -- direct object, 3 -- indirect object, 8 -- chomeur.

2. Clitic pronouns fuse with the preceding element. The examples in this paper give the underlying forms of the verbs, prepositions, and pronouns, rather than the fused forms. The fusion of clitic pronouns with verbs is different from that with **la̲**/**na̲** only in the second person singular. The following chart gives the phonetic forms of the preposition **la̲**/**na̲** followed by the clitic pronouns.

preposition	pronoun	person	phonetic form
la̲	Ø	3	la̲
la̲	i	2s	le
na̲	a	1s	na̲
na̲	á	1pi	ná
na̲	hẽ̲	1pe	na̲hẽ̲
na̲	o	2p	no

While we take these pronominal forms to be (clitic) pronouns throughout this paper because they occur following three major heads (verbs, nouns, and prepositions), an alternative would be to claim that they are agreement affixes rather than pronouns. Just what empirical difference(s) these alternative analyses might make elsewhere in the grammar is not clear to us.

3. It is true for SJM, but apparently not for all other Mazatec languages, that the subject pronoun is sometimes omitted when there is preposition incorporation. We are not able to fully discuss these facts at this time.

4. Apparently it is optional when the 2 is of a certain type, as in the following examples:

(i) miki ?nʧé (-la̲ -Ø) kó co ʧiliö
not hear/1s,3 how say/1s,3
'I don't hear what Julio is saying'

(ii) miki ?nʧé (-la̲ -Ø) so ʃi ti-one
not hear/1s,3 song REL CNT-sound
'I don't hear the song that is playing'

5. An initial 1 is not necessary with the verb 'want', and so the verb also occurs in a structure without Inversion.

(iii) mehẽ̲ sa niño̲ hndé
want more tortillas here
'more tortillas are wanted here'

References

Berinstein, Ava. 1986. Brother-in-law Agreement in Impersonal Constructions. Paper presented at the Conference on Relational Grammar and Grammatical Relations, Ohio State Universtiy.

Bickford, Albert. 1986. The Entrant Term Law. Paper presented at the Conference on Relational Grammar and Grammatical Relations, Ohio State University.

Johnson, David, and Paul Postal. 1980. Arc Pair Grammar. Princeton, New Jersey: Princeton University Press.

Perlmutter, David M., ed. 1983a. Studies in Relational Grammar 1.

----------. 1983b. Personal vs. Impersonal Constructions. Natural Language and Linguistic Theory 1:141-200.

----------. 1983c. The Active Dummy Law. Paper presented at the Grammatical Relations Festival, Cornell University.

----------. 1986. The Overrun Relation in Syntactic Theory: Southern Tiwa Evidence. Paper presented at the Conference on Relational Grammar and Grammatical Relations, Ohio State University.

---------- and Carol Rosen, eds. 1984. Studies in Relational Grammar 2. Chicago: University of Chicago Press.

Postal, Paul M. 1986. Studies of Passive Clauses. Albany: State University of New York Press.

GRAMMATICAL RELATIONS IN ESKIMO: A response to Kalmár

Richard Nivens

1 Introduction
2 Antipassive and the RG analysis
3 On Kalmár's arguments against the relational grammar analysis
4 Other constructions in Iñupiat
  4.1 Passive
  4.2 Indirect object advancement
  4.3 Benefactive-2 advancement
  4.4 Comitative-2 advancement
5 On Kalmár's analysis of Eskimo
  5.1 Case marking
  5.2 Verb agreement
  5.3 Distribution of ergative and antipassive clauses in discourse
6 Conclusion
Notes
References

1 Introduction[1]

Kalmár (1979) attempts to disprove what he sees as two basic assumptions behind the relational grammar analysis of Eskimo antipassives.[2] His arguments, however, have at least two flaws: a misunderstanding of relational grammar, and a less-than-comprehensive view of the Eskimo data.

2 Antipassive and the RG analysis

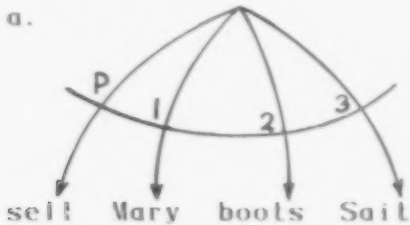
Traditional discussions of Eskimo grammar say that there are two kinds of transitive clause, one called "ergative", the other "antipassive". Seiler and Frantz (to appear) provide the following examples of ergative and antipassive clauses in Iñupiat Eskimo:[3]

(1) a. Mari-m taapkua kamŋ-ich tuni-gai (Saityak-mun)
-ERG those boot-PL sold-3s/3p -F3
'Mary sold those boots (to Saityuk)'

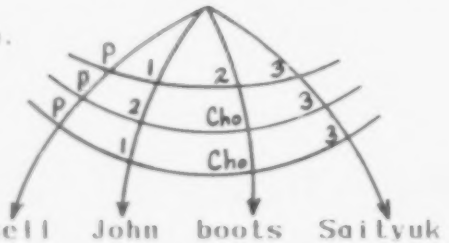
b. Mari-Ø kamiŋ-ñik tunisi-ruq (Saityak-mun)
-ABS -2Cho sell(AP)-3s -F3
'Mary sold boots (to Saityuk)'

The stratal diagrams below indicate the grammatical relations of these two clauses:

(2) a.



b.



The controversial tri-stratal analysis of the antipassive construction was first proposed by Postal (1977). This structure is brought about by interaction between 1-2 Retreat (from a transitive stratum) and the Final 1 Law. The Final 1 Law says that every basic clause must have a 1-arc in the final stratum. Given the three term relations 1, 2, and 3, there are three types of retreat possible: 1-2, 2-3, and 1-3. If the 1 retreats to a 2, placing the initial 2 **en chomage**, the Final 1 Law dictates that something must bear the grammatical relation "1" in the final stratum.

As will be seen, the RG account of antipassive fits in well with a comprehensive generalization of case marking in Iñupiat Eskimo.

3 On Kalmár's arguments against the relational grammar analysis

Kalmár's claim is that RG assumes "that the antipassive clause is generated by a rule whose input is its near-paraphrase, the so-called ergative clause" (p.117). (Later - on page 128 - he makes this just a bit more palatable by saying that the input is the **grammatical relations** of the ergative clause.) He says that RG calls the ergative construction "basic". But in RG, the only thing that is in any way "basic" is the initial stratum of grammatical relations. The ergative construction is not "basic", but it is simpler, having fewer strata than the antipassive -- in fact, only one: the initial stratum is the final stratum. Kalmár is trying to force RG to make claims it doesn't make, by artificially classifying it as a theory that "systematically selects one agnate as 'basic'" (p.128).

Kalmár's lack of understanding of RG naturally causes him to misunderstand the RG analysis of antipassive in Eskimo. For RG does not isolate certain constructions from the others in a language, as Kalmár has done. Nor does it

isolate one language from all other natural languages. Rather, the main goal of relational grammar as a theory is to explain how languages are alike and how they differ.

Kalmár (p.136) cites evidence from Czech to prove that the Eskimo antipassive construction is more similar to the common European transitive clause than the ergative is. He concludes (a) that if either construction in Eskimo is more "basic", it's the antipassive, since the transitive clause is clearly "basic" to European languages; (b) that "subject case-marking and verb agreement are not in themselves sufficient to identify a clause as intransitive". As a result, he claims, the RG arguments against the 2-hood of the noun marked by **-mik/-ñik** are invalid. His logic is that since in Czech the transitive subject is marked the same way as the intransitive subject, and since the verb in both transitive and intransitive clauses agrees only with the subject, then the transitive clause in Czech is more similar to the Eskimo antipassive than to the ergative. However, his argument has a few fundamental flaws. For one thing, no one has ever claimed that Czech had ergative case marking; it is pointless to compare Czech case marking with that of Eskimo. Secondly, the Czech verb never agrees with the direct object, while Eskimo verbs always agree with the (final) direct object, if there is one. His whole argument falls apart. In effect, he accuses RG proponents of doing something they have always argued against: proposing universals based on surface structure.

The RG analysis claims that the antipassive construction is registered by the verbal suffix **-si**.[4] Kalmár doesn't give an argument against this, but he does say in a footnote that **-si** probably "functions to prevent a reflexive interpretation that may attach to semantically transitive verbs without an overt object." But **-si** does occur in clauses with an overt direct object, as was seen in (1b). In such clauses there is no chance of ambiguity. In addition, he offers no explanation for suffixes **-kau** ('PSV'), **-qatigi** ('C2A'), **-uti** ('B2A'), and **-chi** ('32A'), which occur in the same position class as **-si** (see section 3).

Kalmár claims that the intermediate stratum in the antipassive construction "is irrelevant here, although it is important to relational theory for reasons that have little to do with the antipassive clause" (p.130). But the intermediate stratum has much to do with the antipassive clause, as Davies (1984) shows. Davies provides evidence that in Choctaw the verb agrees with the (initial and final) subject of an antipassive clause as if it were a direct

object at some level. In fact, he provides evidence for the grammatical relations of **all three** strata in the Choctaw antipassive construction. The theoretical significance of Davies' paper is that the intermediate stratum **is** relevant, even in languages where it's not explicit that the initial and final subject is an intermediate direct object. Since RG is a theory of universal grammar, evidence from one language can be used to support analyses for all other languages. The tri-stratal RG analysis of antipassive is not a "gimmick", as some have charged: it is a logical result of interaction between various laws of universal grammar which have already proven themselves in countless constructions of numerous languages.

4 Other constructions in Iñupiat

The generalizations that the RG analysis provides are supported by other constructions found in Eskimo, which Kalmár makes no mention of. In this section I will mention only four such constructions, and in section 4 the two analyses will be examined in the light of these constructions.

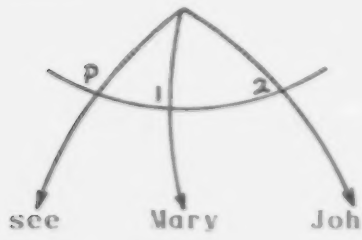
4.1 Passive

The informal RG definition of Passive is "2-1 advancement from a transitive stratum." This construction does occur in Iñupiat, and is registered by the verbal suffix **-kau**.

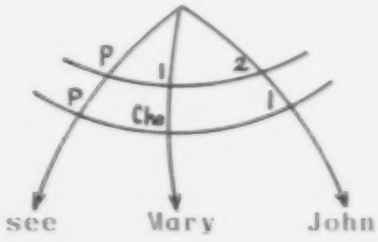
(3) a. Mari-m John-Ø qiñiġ-aa
-ERG John-ABS see-3s/3p
'Mary sees John'

b. John-Ø qiñi-kau-ruq Mari-miñ
-ABS -PSV-3s -1Cho
(same)

(4) a.



b.



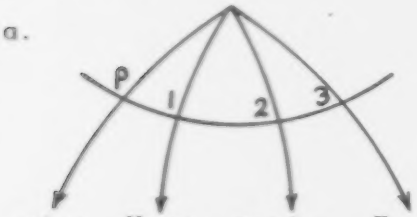
4.2 Indirect object advancement

Many languages allow an initial indirect object to advance to direct objecthood, thus placing the initial direct object (if any) **en chomage**. In Iñupiat this advancement is registered by the verbal affix -chi. Consider the following pair of sentences:

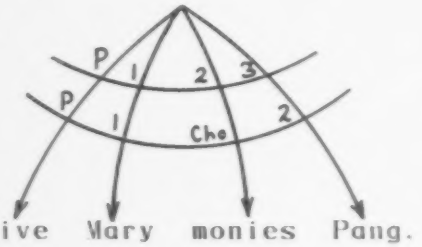
(5) a. Mari-m mani-ich qait-kai Paŋaliŋ-mun
-ERG money-PL give-3s/3p -F3
'Mary gave the money to Pangalik'

b. Mari-m Paŋalik-Ø qait-chi-gaa maniŋ-ñik
-ERG -ABS -32A-3s/3s money-2Cho
(same)

(6) a.



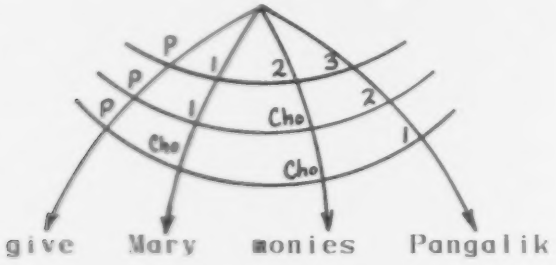
b.



Consider also the following sentence and its corresponding stratal diagram:

(7) Paŋalik-Ø qait-chi-kau-ruq maniŋ-ñik Mari-miñ
-ABS -32A-PSV-3s money-2Cho -1Cho
'Mary gave the monies to Pangalik'

(8)



Since the initial 3 has advanced to 2, it is eligible to advance to 1 via Passive. The initial 2, on the other hand, cannot advance:

(9) a.* mani-ich Paŋalik-Ø qait-chi-kau-rut Mari-miñ
-3p

b.* mani-ich Paŋalik-Ø qait-chi-kau-gaat Mari-miñ
-3p/3s

The fact that the initial 2 cannot advance (via Passive) can be taken as evidence that the initial 3 has placed it **en chomage**.

4.3 Benefactive-2 advancement

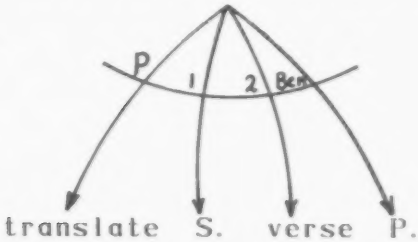
Various obliques can also advance to termhood: one such possibility is Benefactive-2 advancement. Consider the sentences in (10), and note the presence of suffix **-uti**:

(10) a. Siqupsira-m taiyuaq-0 mumik-kaa Paŋaliŋ-mun
-ERG verse-ABS translate-3s/3s -BEN
'Siqupsiraq translates a verse for Pangalik'

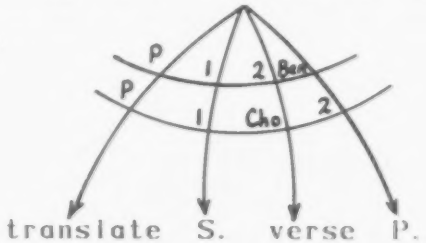
b. Siqupsira-m Paŋalik-0 mumi-uti-gaa taiyua-mik
-ERG -ABS -B2A-3s/3s -2Cho
(same)

Seiler (1978:74) proposes the following relational networks:

(11) a.



b.



4.4 Comitative-2 advancement

Another oblique that can advance to direct object is the comitative. In fact, Comitative-2 advancement is obligatory in Iñupiat, whether the initial stratum is intransitive as in (12) or transitive as in (13):

(12) a. Putu-0 aullaq-tuq
-ABS leave-3s
'Putu went away.'

b. Putu-m Matulik-0 aullaq-qatigi-gaa
-ERG -ABS leave-C2A-3s/3s
'Putu went away together with Matulik.'

(13) a. Mari-m kuvraq-0 amu-gaa
-ERG net-ABS pull:out-3s/3s
'Mary pulls out the net.'

b. Mari-m John-Ø amu-qatigi-gaa kuvra-mik
-ERG -ABS pull:out-C2A-3s/3s net-2Cho
'Mary together with John pulls out the net.'

This obligatory advancement is registered on the verb with the suffix **-qatigi**.

5 On Kalmár's analysis of Eskimo

"Ergative and antipassive have the same grammatical relations" (p.117). This statement of Kalmár's is both true (for the initial stratum) and false (for the structure as a whole). A comprehensive view of facts concerning verb agreement and case-marking generalizations will bear this out.

5.1 Case marking

There is a problem with Kalmár's notions of "subject" and "direct object". He defines "subject" syntactically, relying on Keenan's list of subject properties -- a list shown by Dixon (1979:110-12) to be weak or even invalid for so-called "ergative languages" like Eskimo. But even if we grant the validity of Kalmár's argument concerning subjects, he still has to prove that the noun marked with **-mik** is a direct object. Identifying one nominal as a subject does not make the other a direct object. Indeed, Kalmár himself is hesitant to call the **mik** noun a direct object (p. 120), saying he will deal with that question later. Yet throughout the article he does refer to it as a direct object: I find no argument to the contrary. Actually, he seems to be basing his notion of "direct object" on semantic criteria. He seems to define "direct object" as the patient in a clause, or (more likely) as what Dixon (1979) calls the "O" nominal. In other words, he uses the term "direct object" for initial direct objects, while he defines "subject" on the basis of surface structure criteria.

Relational grammar provides a straightforward accounting of case-marking in Iñupiat:

(14) Final grammatical relations are marked as follows:
-m marks singular final ergative nominals
-Ø marks singular final absolutive nominals
-ich marks plural final (nuclear?) terms
-mik (PL **-ñik**) marks final 2-chomeurs and instruments[5]
-miñ (PL **-niñ**) marks final 1-chomeurs and sources
-mun (PL **-nun**) marks final 3s, benefactives, and goals

Kalmár, on the other hand, sees the ergative/antipassive distinction as "basic", and his description of case marking is split along this line:

(15) a. In ergative clauses:
-up marks subjects
-Ø marks direct objects

b. In antipassive clauses:
-Ø marks subjects
-mik/-ñik marks direct objects

c. In intransitive clauses:
-Ø marks subjects

To handle the other constructions I cited in section 3, Kalmár's description of case marking would be very complex:

(16)	Subj.	D.O.	I.O.	BEN	COM
ergative (1a,3a,5a,11a)	-m	-Ø	-mun	-mun	--
passivized ergative (4b)	-miñ	-Ø	-mun	-mun	--
32A clause (5b)	-m	-mik	-Ø	-mun	--
passivized 32A (7)	-miñ	-mik	-Ø	-mun	--
B2A clause (10b)	-m	-mik	-mun	-Ø	--
antipassive (1b)	-Ø	-mik	-mun	-mun	--
intransitive	-Ø	--	--	-mun	--
intransitive with COM (12b)	-m	--	--	-mun	-Ø
transitive with COM (13b)	-m	-mik	-mun	-mun	-Ø

5.2 Verb agreement

With regard to verb agreement, Kalmár again misses the generalization that RG provides by insisting on a basic dichotomy of ergative and antipassive clauses. The generalization for verb agreement is really quite simple:

(17) The verb agrees in person and number with final nuclear terms.

But Kalmár's description of verb agreement would contain no generalization. In simple antipassive clauses there is no final direct object and therefore no person agreement with direct object.

(18)

clause type:	verb agrees with:
ergative	- subject and direct object
passivized ergative	- direct object
32A clause	- subject and indirect object
passivized 32A	- indirect object
B2A clause	- subject and benefactive
antipassive	- subject
intransitive	- subject
intransitive with COM	- subject and comitative
transitive with COM	- subject and comitative

5.3 Distribution of ergative and antipassive clauses in discourse

Kalmár proposes that the antipassive structure is used when the direct object is "new", while the ergative structure is used for "given" direct objects. He claims that -■ik causes a new referent to be entered into the hearer's "registry", and that it is also a morphological trigger which causes the syntax of the sentence to change. The former claim I can accept, but the latter doesn't do justice to the data. If the presence of -■ik is sufficient to mark a new referent, why should any change in syntax occur at all? If the antipassive subject is still the subject of a transitive clause, as Kalmár claims, why should it be marked with the Ø case, which happens to be the case marking for intransitive subjects? Case marking in Eskimo is clearly on an ergative/absolutive basis. Why is the verbal suffix -si found only in antipassive constructions? Why should the verb cease to agree with the direct object, just because it's "new"? Why should the syntax change when there is no overt morphological trigger, i.e. when the direct object is unspecified? And why only direct objects? If -■ik indicates that a referent is to be entered into the hearer's "registry", why aren't new subjects marked with -■ik? Kalmár's "explanation" raises more questions than it answers.

In fact, it just doesn't work. He identifies the "new" direct object with the -■ik case. But he only discusses the -■ik case in antipassive clauses. Recall that the -■ik case also occurs in clauses with 3-2 advancement (5b, 7), in clauses with Benefactive-2 advancement (10b), and in clauses with Comitative-2 advancement (13b). While -■ik may well mark "new" direct objects in antipassive clauses, it seems unlikely that it would mark "new" direct objects in clauses involving the promotion of another nominal. Thus "new" direct objects might better be identified with the antipassive clause, rather than with the -■ik case.

Kalmár (p.124) states that antipassive clauses are more common in elicited material than are ergative, while the reverse is true of discourse. He explains that since elicited sentences are generally devoid of a discourse context, it makes sense that the information encoded in them should tend to be "new". According to the relational grammar analysis, the ergative construction is syntactically simpler than the antipassive. Since "given" information is what holds a discourse together, we might expect the simpler (shall we say "unmarked"?) construction to be used for this "basic" discourse function, and thus be more common in narrative texts, as is the case. The more complex construction should be used for special purposes, such as entering a new referent into the hearer's "registry". But for some reason Kalmár doesn't see it this way. He doesn't see the introduction of new information as any more specialized a function than the maintenance of given information. Rather, he sees these two discourse functions as "symmetrical", "diametrically opposed", "equivalent" (p.133). It is for this reason - because of his view of discourse functions - that he insists that the ergative construction is syntactically no simpler than the antipassive.

Relational grammar, like TG, is a theory of "autonomous syntax". It does not claim that other factors (semantic, pragmatic, thematic, etc.) are irrelevant; however, it claims that syntax itself can be adequately described without reference to these factors. Perlmutter (1980:203-4) states that

> a particular construction may be linked in individual languages with semantic, pragmatic, or presuppositional effects, with constraints on definiteness or specificity or reference of nominals, with the organization of the sentence into old and new information, and so on. The general strategy of RG in all such cases is to separate the syntactic nature of a particular construction from the semantic, pragmatic, etc., factors with which it interacts. ... A particular syntactic construction can be the same in two languages that use it in very different semantic or pragmatic contexts.

6 Conclusion

An analysis which seems adequate for simple constructions in a language may prove to be inadequate when more complex constructions are considered. Kalmár's

analysis of antipassive in Eskimo becomes burdensome when you look at its implications for a comprehensive analysis of all clause types. Relational grammar gives a much simpler account of case marking and verb agreement in Eskimo. The examples given here should prove the point; Seiler (1978) and Seiler and Frantz (to appear) also provide evidence from causative clause union, relative clauses, participial groups, and reflexives. And not only does RG provide a solid language-internal analysis, it also shows how Eskimo is similar to other languages, and how it is different.

Notes

1. I am grateful to Chuck Speck, Cindy Williams, Steve Marlett, Dave Weber, Don Frantz, and John Little for comments on earlier drafts of this paper. Of course, I take responsibility for any errors or inadequacies. For a brief introduction to relational grammar, see Perlmutter (1980).

2. "The first is that the antipassive clause is generated by a rule whose input is its near-paraphrase, the so-called ergative clause. The second assumption is that the antipassive clause has no superficial direct object, and is therefore a "surface" intransitive clause" (p.117).

3. Data not taken from Kalmár's paper are from Seiler (1978) or Seiler and Frantz (to appear). (Kalmár's -up and -it correspond to the Iñupiat -**ŋ** and -ich.) Abbreviations used here are: ABS = absolutive; AP = antipassive; BEN = benefactive; B2A = BEN-2 advancement; Cho = chomeur; COM = comitative; C2A = COM-2 advancement; ERG = ergative; F3 = final indirect object; PL = plural; PSV = passive; 1 = subject; 2 = direct object; 3 = indirect object; 32A = 3-2 advancement. In verb glosses only: p = plural; s = singular; 3 = 3rd person; 3s/3p = 3s final 1, 3p final 2.

4. The suffix -si has several allomorphs. Woodbury (1977:323) reports that in Greenlandic Eskimo, -si has allomorphs -i, -si, -ši, -ŋŋig, -llir, or -Ø, depending on the verb stem.

5. RG currently makes no attempt to explain **why** various grammatical relations share the same case marking. This is not to say, however, that no cross-linguistic generalizations can be found.

References

Davies, William D. 1984. Antipassive: Choctaw evidence for a universal characterization. In Studies in relational grammar 2, ed. David Perlmutter and Carol Rosen. Chicago: University of Chicago Press. pp.331-76

Dixon, R.M.W. 1979. Ergativity. Language 55.1.59-138.

Kalmár, I. 1979. The antipassive and grammatical relations in Eskimo. In Ergativity: Toward a theory of grammatical relations, ed. Frans Plank. New York: Academic Press. pp.117-143.

Perlmutter, David. 1980. Relational grammar. In Syntax and semantics 13: Current approaches to syntax, ed. E.A. Moravcsik and J.R. Wirth. New York: Academic Press. pp.195-229.

Postal, Paul. 1977. "Antipassive in French." In Proceedings of the Seventh Annual Meeting of the North East Linguistic Society. Cambridge, Mass.:MIT. Reprinted in Lingvisticae Investigationes 1:333-74.

Seiler, Wolf. 1978. The modalis case in Iñupiat. In 1978 work papers of the Summer Institute of Linguistics, University of North Dakota, vol. 22, ed. John P. Daly. Huntington Beach, CA: SIL. pp.71-85.

Seiler, Wolf, and Don Frantz. To appear. The instrumental case in Iñupiat (Eskimo).

Woodbury, A.C. 1977. Greenlandic Eskimo, ergativity, and RG. In Syntax and semantics 8: Grammatical relations, ed. P. Cole and J.M. Sadock. New York: Academic Press. pp.307-36.

A RELATIONAL GRAMMAR APPROACH TO VERB AGREEMENT IN LAKOTA

Gray Plunkett and Michael McKeever

1 Introduction
2 Relational Grammar
3 Basic verb agreement in Lakota
4 Simple monostratal transitive clauses in Lakota
5 Agreement in intransitive clauses
6 Reflexive clauses
7 Double patient verbs
8 3-2 and Obl-2 advancements
9 The notion of working 2
10 Conclusion
Notes
References

1 Introduction

This paper examines verb agreement rules and their application in Lakota, a Siouan language.[1] In particular, it explores how to best account for verb agreement in Lakota in a Relational Grammar framework. We will argue that certain concepts of Relational Grammar (RG), especially the notions of level and working 2, provide the necessary theoretical apparatus for formulating concise rules of verb agreement in Lakota. These rules account for all major transitive, ditransitive, and intransitive clause types. It also lends support to Perlmutter's claim that the notion of working term-x is needed to state rules in the grammars of natural language (Perlmutter 1982). We will also show the importance of disjunctive ordering in the grammars of natural languages.

In Sect. 2 of this paper we will give a brief introduction to Relational Grammar. In Sect. 3 we will show the two sets of verbal agreement markings and how they register on the verb. In Sect. 4 we consider simple monostratal transitive clauses in Lakota and present an initial hypothesis concerning their agreement. In Sect. 5 we present intransitive clauses and their effect on our initial hypothesis. Sect. 6 deals with reflexive clauses. In Sect. 7 we discuss 'double patient' verbs in Lakota and in Sect. 8 3-2 and Obl-2 Advancements. Sect. 9 introduces the notion of working 2 and its importance in formulating

the final verb agreement rule for Lakota. Conclusions and implications are discussed in Sect. 10.

2 Relational Grammar

One of the primary goals of Relational Grammar is to construct adequate and insightful grammars for individual languages. A basic contention of Relational Grammar is that semantic roles such as Agent, Experiencer, Patient etc. are insufficient for this purpose. Instead, Relational Grammar claims that to do this primitive grammatical relations, such as 'subject of,' 'object of,' 'indirect object of,' etc. are needed. Although these are strictly primitives, the Universal Alignment Hypothesis does state that there is a loose connection between grammatical relations at an initial level (discussed below) and semantic roles (Perlmutter and Postal 1984:97). (It is at this point that semantics and syntax are related in RG.)

The following is a partial list of grammatical relations used in RG.

TERMS	ABBREVIATIONS
subject	1
direct object	2
indirect object	3
NONTERMS	
OBLIQUE	
Benefactive ("for")	Ben
Directional ("to")	Dir
Locative ("at", "in")	Loc
RETIREMENT	
Chomeur	Cho

The oblique relations closely correspond to the semantic roles of the same name. The Chomeur is the relation that a nominal bears when another nominal assumes its relation.

Rosen (1984:40) in discussing semantic roles and grammatical relations quotes a form of Perlmutter and Postal's Universal Alignment Hypothesis which states:

> There exists some set of universal principles on the basis of which, given the semantic representation of a clause, one can predict which initial GR each nominal bears.

Although Rosen argues against the above hypothesis as a universal in its present form, for the purpose of illustration let us show what some tentative predictions from this hypothesis might be.

(1) Agents, experiencers, and cognizers are (initial) 1s
Patients and stimuli are (initial) 2s.
Recipients and addressees are (initial) 3s.
Beneficiaries are (initial) Benefactives.
Instruments are (initial) Instrumentals.

In a sentence such as:

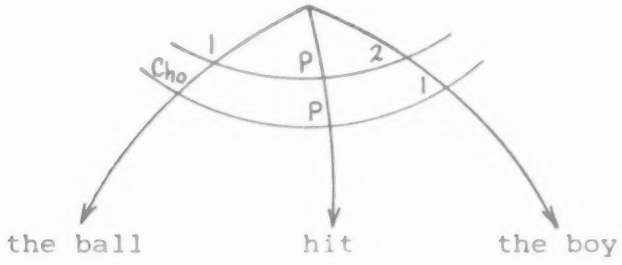
(2) The boy hit the ball.

The agent 'the boy' is the subject (or 1), the patient 'the ball' is a direct object (or a 2). These facts are correctly predicted by the principles in (1), yet in a sentence such as:

(3) The ball was hit by the boy.

the agent is not the subject and the patient is not the direct object. This sentence seems to counter the principles in (1). Yet this is not the case once a second important concept of Relational Grammar, the notion of levels, is taken into consideration. In RG a passive sentence such as (3) is analyzed as having two levels of syntactic structure, an initial level and a final level. In a stratal diagram sentence (3) would look like:

(4)



The diagram above says that 'the boy' heads a 1-arc in the initial level but in the final level it heads a chomeur (Cho). 'The ball' heads a 2-arc initially but a 1-arc finally. The verb 'hit' heads a P or predicate arc. (The order of constituents in a stratal diagram is inconsequential.) Note that this analysis does not violate the principles stated in (1). The stratal diagram in (4) represents a static relationship as opposed to the movements

posited for a passive transformation in a Transformational framework. By referring to grammatical relations at different levels, one can easily capture language particular generalizations about word order, pronominal case, verb agreement, etc. Thus, Relational Grammar employing grammatical relations and multiple syntactic levels has been shown to formulate insightful and adequate grammars in many diverse languages.

3 Basic verb agreement in Lakota

Verbs in Lakota indicate person agreement by prefixes and infixes. (There are also emphatic free standing pronouns. These forms will be used only for convenience to represent the nominals in diagrams.) We see two verbal agreement sets occurring on the verb plus a suffix **-b** to show plural. The subject agreement set is observed in the paradigm in (5) (of the verb **mani** 'walk').[2]

(5) Subject Agreement Set

ma-**wa**-ni	I walk
ma-**ya**-ni	you walk
ma-**ø**-ni	he/she/it walks
ma-**ų**-ni	we (inc.) walk
ma-**ų**-ni-**b**	we (exc.) walk
ma-**ya**-ni-**b**	you (pl.) walk
ma-**ø**-ni-**b**	they walk

Object agreement is observed in the paradigm in (6) (on the verb **apa** 'hit'):

(6)

a-**ø**-**ma**-**pa**	he hit me.
a-**ø**-**ni**-**pa**	he hit you.
a-**ø**-**ø**-**pa**	he hit him.
a-**ø**-**ų**-**pa**-**b**[3]	he hit us (inc. & exc.).
a-**ø**-**ni**-**pa**-**b**	he hit you (pl.).
a-**ø**-**wi**-**pa**	he hit them.

We can now list the two agreement sets.

(7)

	Set I Subject	Set II Object
1S	**wa-**	**ma-**
2S	**ya-**	**ni-**
3S	**ø**	**ø**
1P	**ų-**	**ų-**
2P	**ya-**	**ni-**
3P	**ø**	**wiđa-**

Several points should be made concerning the preceding paradigms. First, some verbs infix the agreement markers, while other verbs prefix agreement markers (ex. **wa-loǫ** 'I sing'). However, many cases of apparent infixing could be due to the presence of locative or verbal prefixes which precede agreement markers. Lakota has an extensive inventory of verbal prefixes requiring further study. These prefixes often trigger morphophonemic processes which can alter the agreement affixes.[4] Such processes are beyond the scope of this paper. Second, (5) shows that Lakota has an inclusive-exclusive distinction in the 1st person plural subject marked by the absence or presence of the plural morpheme **-b**. Third, the 3rd person plural marker **wiđa-** only appears with transitive verbs and only marks animate 3rd plural objects. Its use is slightly different from the rest of the Set II affixes.

4 Simple monostratal transitive clauses in Lakota

A look at simple transitive clauses reveals that Set I and II are used for subject and direct object agreement respectively. Example (8) shows this agreement.

(8) a-**ma-ya-**ƥe
hit-1S:II-2S:I-hit
'You hit me.'

When **wa-** and **ni-** would be expected together, a portmanteau morpheme **đi-** is used instead as in (9).

(9) a-**đi-**ƥe
hit-1:2-hit
/a+**wa+ni+**ƥe/
'I hit you.'

The plural marker **-b** appears with all plural subjects (except for 1 plural inclusive) as well as with plural objects of 1st and 2nd person.

(10) he a-**ya**-**ƥa**-b
PRO:3S hit-2S:I-hit-PL
'You(pl.) hit him.'

(11) he a-**ni**-**ƥa**-**b**
PRO:3S hit-2P:II-hit-PL
'He hit you (pl.).'

wiƈa- appears with all animate 3rd person plural objects.

(12) hena a-**wiƈa**-**wa**-**ƥe**
PRO:3P hit-3P:II-1S:I-hit
'I hit them.'

(13) hena a-**wiƈa**-**ya**-**ƥa**-**b**
PRO:3P hit-3P:II-1S:I-hit-PL
'You(pl.) hit them.'

The order of the morphemes is significant and has the following relative order: **ų-, ma-, wiƈa-, ya-, wa-, ni-**. Examples of the application of this ordering are in (8), (12), (13) and the following examples.

(14) **ųk**-a-**ni**-**ƥa**-**b**
1P:I-hit-2:II-hit-PL
'We hit you (s. or pl.).'

(15) hena **ųk**-a-**wiƈa**-**ƥa**-**b**
PRO:3P 1P:I-hit-3P:II-hit-PL
'We hit them.'

Based on the above examples, verb agreement in Lakota could be expressed by the following statements:

(16) Nominals heading a 1-arc determine Set I agreement markers.

Nominals heading a 2-arc determine Set II agreement markers.

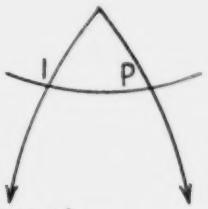
Since all the examples considered here involve only one syntactic level it is not possible to tell if the agreement is based on the initial or final level. In order to do this we will turn to a consideration of intransitive verbs.

5 Agreement in intransitive clauses

Within the framework of Relational Grammar it has been claimed that there are two kinds of intransitive verbs--those with an initial 1 (called unergative) and those

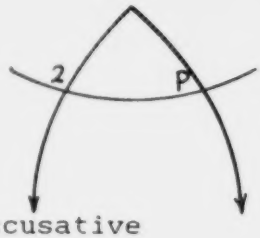
with an initial 2 but no initial 1 (called unaccusative). The initial strata for these verbs are represented in the following stratal diagrams in (17):

(17) a.



unergative

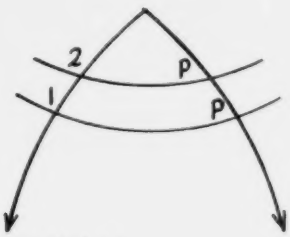
b.



unaccusative

The Final 1 Law proposed by Perlmutter and Postal (1983:100) posits that (17b) must have a 2-1 Advancement resulting in the following stratal diagram in (18):

(18)



unaccusative

Perlmutter and Postal's analysis of unaccusative verbs provides the theoretical apparatus one needs to state the verb agreement rules with greatest generality. In this section we will present how this analysis accounts for intransitive clauses in Lakota and how this information requires a further refinement of the verb agreement rule in (16).

A salient fact of Lakota is the existence of two groups of intransitive verbs which show different verb morphology. Perlmutter and Postal (1984:98-100) point out that Boas and Deloria's (1941) description of intransitive verbs (which take either Set I or Set II agreement markers) corresponds to their proposed categories of unergative and unaccusative verbs.[5] The following data illustrates the two types of intransitive verbs in Lakota:

(19) **ya**-psiƌa-**b**
2P:I-jump-PL
'You (pl.) jump.'

(20) **ma**-cuwita
1S:II-cold
'I'm cold.'

The verb in (19) takes a Set I agreement marker while the verb in (20) takes Set II. Based on semantics one would expect 'jump' to be an ergative verb and 'cold' to be unaccusative. According to rule (16), the unergative clause should take Set I agreement markers while the unaccusative clause should take Set II agreement markers. This fits the data in (19) and (20).

However, once the analysis presented in (18) for unaccusative verbs is adopted, the verb agreement rule stated in (16) is no longer explicit. The argument to the predicate in (20) bears both the subject and direct object relation to the clause but at different levels. Therefore the verb agreement rule must be modified to refer to level. The rule could be restated as follows:

(21) Initial 1s determine Set I agreement markers.

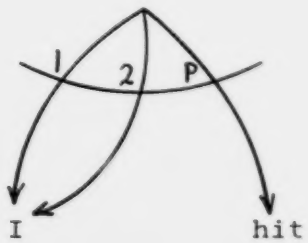
Initial 2s determine Set II agreement markers.

This generalization accounts for all the transitive and intransitive clauses considered so far.

6 Reflexive clauses

We will now consider a reflexive clause. Within the framework of RG, reflexive clauses are considered to be a case of multi-attachment. This means that at the initial level the single nominal heads both a 1- and a 2-arc as diagramed below.

(22)



'I hit myself'

Williamson (1979:357-359) has given evidence from restrictions on reflexivization in Lakota that the nominal

in a reflexive clause does head both a 1- and a 2-arc at the initial level. In Lakota the verbal affix **iȼ'i-** indicates a reflexive. Consider now the reflexive clauses in (23) below.

(23) a. a-**m**-iȼ'i-pa
hit-1S:II-REFL-hit
'I hit myself.'

b. a-**n**-iȼ'i-pa
hit-2S:II-REFL-hit
'You hit yourself.'

It can be seen from the diagram in (22) that reflexives such as those in (23) have a nominal heading both an initial 1- and 2-arc. The verb agreement hypothesis as stated in (21) has no way of predicting whether one or both of the arcs will determine agreement for the nominal of a reflexive. We must have a way of specifying in our rule which arc is determining agreement.

We can see from the clauses in (23) that it is in fact only the 2-arc that is determining verbal agreement. To write a verbal agreement rule which handles reflexive constructions we must take into consideration disjunctive ordering in which, if a nominal fulfills two rules, only one applies.

In proposing disjunctive ordering of agreement rules we will follow Davies (1981:306) in saying that they are applied the same as phonological rules. For example, the French stress rule in (24) shows two disjunctively ordered rules.

(24) a. $V \rightarrow [+\text{stress}] / ____ \; C_0 \; \underset{[-\text{tense}]}{V} \; \#$

b. $V \rightarrow [+\text{stress}] / ____ \; C_0 \; \#$

The rules are disjunctively ordered so that in case a single form satisfies both parts, only (24a) applies. This disjunctive ordering prevents ungrammatical forms from being generated. This is the type of ordering being proposed for Lakota agreement rules. The new verb agreement rule is now:

(25) a. Initial 2s determine Set II agreement markers.

b. Initial 1s determine Set I agreement markers.

Where (a) is disjunctively ordered with respect to (b).

7 Double patient verbs

There is a small set of verbs in Lakota which take two Set II agreement markers. Boas and Deloria (1941:76-77) call these neutral (stative) verbs with two objects. They state that certain neutral verbs implying comparison may take two object agreement markers. Williamson (1979:359) refers to these as 'Double Patient' verbs. The following shows six examples which take 'Double Patients.'[6]

(26) i-**ma-ni-**ǧteč̌e
ashamed-1S:II-2S:II-ashamed
'You are ashamed of me.'

(27) iye-**ni-ma-**č̌eca
resemble-2S:II-1S:II-resemble
'I resemble you.'

(28) hena iye-**wiča-ma-**č̌eca
PRO:3P resemble-3P:II-1S:II-resemble
'I resemble them.'

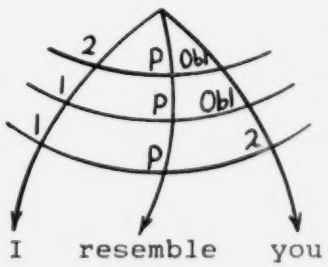
(29) iyo-**ni-ma-**kpi
happy-2S:II-1S:II-happy
'I am happy about you.'

(30) i-**ni-ma-**skola
small-2S:II-1S:II-small
'I am as small as you.'

An analysis of these clauses based on the verb agreement rule in (25) would require that both nominals are initial 2s. However, this would be a violation of the Stratal Uniqueness Law of Perlmutter and Postal (1983:92) which claims that two nominals cannot bear the same term relation in a given stratum.

A more plausible analysis for these clauses, based on Williamson (1979:361), is presented below by the stratal diagram in (31).

(31)



Williamson argues, based on the Universal Alignment Hypothesis, that most of the verbs which take this construction are represented in English by verbs which take obliques. Taking into consideration this analysis, a verb agreement rule which would account for 'double patient' clauses as well as those previously considered could be the following:

(32) a. Nominals heading a 2-arc determine Set II agreement markers.

b. Nominals heading a 1-arc determine Set I agreement markers.

Where (a) is disjunctively ordered with respect to (b).

The phrase 'nominals heading an x-arc' refers to all nominals which bear the x-relation at some level. Such a generalization would work for the analysis in (31) since the two nominals each head a 2-arc and thus would take Set II agreement markers. However, a problem arises with this analysis when we consider 3-2 and Oblique-2 Advancements in Lakota.

8 3-2 and Obl-2 advancements

Certain advancements to 2 are obligatory in Lakota. They are shown by the fact that the verb shows agreement with the initial 3 or some initial oblique. The following clauses show clear cases of 3-2 and Obl-2 Advancements in Lakota.

(33) **ʒų**ka **wą** **ma**-k'u
dog INDF 1S:II-give
'He gives me a dog.'

(34) **įyą** ki **ma-ya**-k'u
rock DEF 1S:II-2S:I-give
'You give me the rock.'

(35) he **ʒų**ka **wą** **ų**-kipazo-**b**
PRO:3S dog INDF 1EXC:II-show-PL
'He showed us (exc.) a dog.'

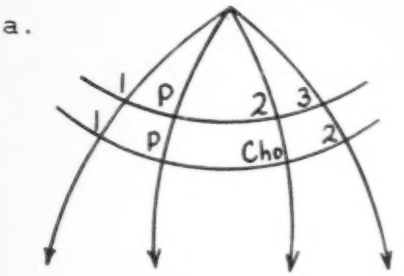
(36) he wo-**ni**-**ʤ**iyake
PRO:3S UO-2S:II-talk
'He talks to you.'

(37) **ʤ̨ka wą** o**p̣e-ʤi-ʤ**ato
dog INDF buy-1:2-buy
'I bought you a dog.'

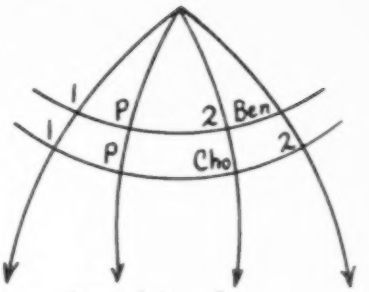
(38) **ʤ̨ka wą** o**p̣e-ma-ya-ʤ**ato
dog INDF buy-1S:II-2S:I-buy
'You bought me a dog.'

The analysis of (34) & (35) would be:

(39) a.



b.



According to rule (32) nominal **a** should determine Set I agreement and nominal **c** Set II. This is exactly what happens in (34) and (35). However, rule (32) would also predict that nominal **b** would show Set II agreement since it also heads a 2-arc. The following clause shows that nominal **b** does not show Set II agreement:

(40) * he **ma-ni**-kipazo
PRO:3S 1S:II-2S:II-show
'He showed you to me.'

Instead the correct form is:

(41) he niye **ma**-kipazo
PRO:3S PRO:2S 1S:II-show
'He showed you to me.'

Examples (40) and (41) show that the 2nd person initial 2 does not show agreement with the verb with Set II markers but instead must be an independent pronoun. (For a discussion of independent pronouns in Lakota see Boas and Deloria (1941:78).) It appears from this example that final nominals which are 2-chomeurs do not determine verb agreement in Lakota. The following example lends support to this hypothesis:

(42) he niye **wa**-kipazo
PRO:3S PRO:2S 1S:I-show
'I showed you to him.'

In (42) 'you' is an initial 2 but a final 2-chomeur. The verb does not show verb agreement with this nominal. Instead the verb is showing agreement with the nominal heading a 1-arc ('me') and the nominal 'he' which is a final 2. Since the final 2 is 3rd person singular in (42) there is no overt marking on the verb for the final 2.

Thus any verb agreement rule must exclude 2-chomeurs. This rules out the use of the phrase 'heading a 2-arc', since all 2-chomeurs are 2s at a previous level.

Another possible category would be acting 2. Perlmutter (1982) defines the notion of acting term on page 307 of his article. The term acting 2, for example, groups together final 2s and final 2-chomeurs. It would, therefore, not be useful in stating verb agreement in Lakota for the same reason as the notion of heading a 2-arc. It would also not work for unaccusative clauses since they do not contain acting 2s.

9 The notion of working 2

Perlmutter (1982:314) defines the notion of working term-x as being a nominal which heads a term-x-arc and bears a term relation in the final stratum. By definition this rules out all 2-chomeurs as working 2s. The notion of working 2 allows the formulation of the following verb agreement rule for Lakota:

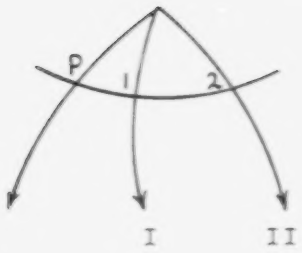
(43) a. Working 2s determine Set II agreement markers.

b. Nominals heading a 1-arc determine Set I agreement markers.

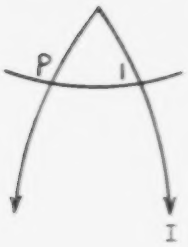
Where (a) is disjunctively ordered with respect to (b).

This rule accounts for the proper agreement markers for all the clause types and analyses proposed so far. The following schematically represents the data presented in this paper:

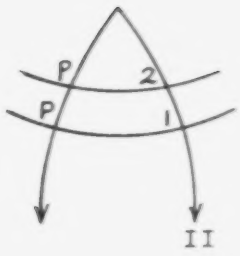
(44) a.



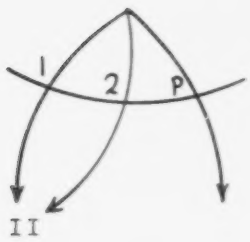
b.



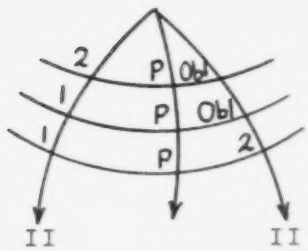
c.



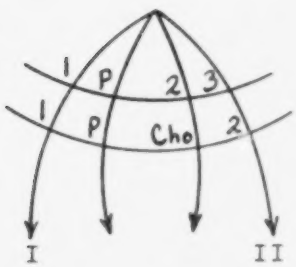
d.



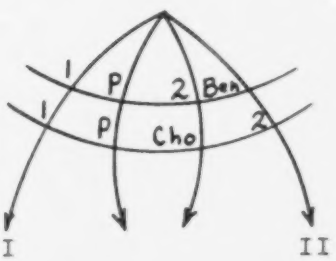
e.



f.



g.



In the above stratal diagrams all working 2s take Set II agreement markers. No other generalization would account for the varied cases where Set II agreement markers appear. There is no evidence in the data presented thus far which specifies at which level Set I agreement markers are determined.

10 Conclusion

Rule (43) is a satisfactory verb agreement rule for Lakota. It accounts for the use of Set I and Set II agreement markers in a wide variety of data and constructions. It accounts for all transitive and intransitive, single level and multilevel, and single attachment and multi-attachment clauses. It also seems to account for seemingly irregular registration such as 'double patient' clauses.

Disjunctive rule application is a topic of current interest in the field of inflectional morphology. Davies (1981:324) has hypothesized that disjunctive rule application of agreement rules is a universal principle. He states, "Given a set of agreement rules making a predicate agree with the same properties a, b....n of nominal, the rules apply disjunctively." It is interesting to note that recent work done by Steve Willson (1985)in Burushaski, a Pakistani language has produced evidence to counter Davies' view of disjunctive application as a universal principle. While disjunctive ordering may not prove to be a universal principle for instances in which a nominal satisfies the requirements of two separate rules, it seems crucial to account for verb agreement in Lakota.

This study has shown the value of the RG concepts of syntactic levels. It has especially given support to Perlmutter's claim that the notion of working term-x should be available in formulating concise rules in the grammar of languages. Further investigation is needed into other syntactic constructions of Lakota, especially clause union, to substantiate the viability of rule (43). However, this is a start at a working generalization for determining verb agreement in Lakota. The framework of Relational Grammar has provided a means for making simple, concise rule statements and analyses for some seemingly complicated agreement data.

Notes

1. Lakota is a dialect of Dakota-Sioux, belonging to the Siouan language family of North and South Dakota, Nebraska, Minnesota and Canada. Dakota-Sioux is spoken by about 25,000 people. The data contained in this paper comes from Walter Taken Alive from Standing Rock Reservation, South Dakota and Bert McBride originally from Crow Creek Reservation, South Dakota during the summer session of the

Summer Institute of Linguistics at the University of North Dakota in 1985.

2. The following abbreviations will be used in this paper: 1S=1st person singular, 2S=2nd person singular, 3S=3rd person singular, 1P=1st person plural, 2P=2nd person plural, 3P=3rd person plural, 1:2=1st person subject:2nd person object, I=Set I, II=Set II, INC=inclusive, EXC=exclusive, DEF=definite article, INDF=indefinite article, PL=plural marker, PRO=pronoun, REFL=reflexive, and UO=unspecified object.

3. A variant of this form is **ųk-a-ƥa-b**. It appears that only **ų/ųk** can show this alteration in the placement of the affix.

4. Another very common morphophonemic change occurs with y-initial verbs in 1st and 2nd person which would appear to be an exception to the two set generalization. There is evidence that only active y-initial verbs that would normally take Set I undergo a morphophonemic change while Set II verbs do not undergo such a change. These and other morphophonemic processes will not be dealt with in this paper.

5. Rosen (1984) argues that Lakota provides evidence that a purely semantic account for the differences between unergative and unaccusative clauses is not adequate.

6. Since both agreement markers are from Set II, the ordering hierarchy of morphemes is replaced by a rule that requires the morpheme showing agreement with the initial subject to go last.

References

Boas, Franz and Deloria, Ella. 1941. Dakota Grammar. Washington: U.S. Government Printing Office.

Davies, William. 1981. Choctaw Clause Structure. Phd. Dissertation. San Diego: University of San Diego.

Perlmutter, David M. 1980. Relational Grammar. In Syntax and Semantics 13: Current Approaches to Syntax. E. Moravcsik and J. Wirth, eds. New York: Academic Press.

----------. 1982. Syntactic representations, syntactic levels, and the notion of subject. In The Nature of

Syntactic Representation. P. Jacobson and G. Pullum, eds. Dordrecht and Boston: D. Reidel Publishing Company.

Perlmutter, David M. and Postal, Paul M. 1983. Some proposed laws of basic clause structure. In Studies in Relational Grammar 1. David Perlmutter, ed. Chicago: University of Chicago Press. 81-128.

----------. 1984. The 1-Advancement Exclusiveness Law. In Studies in Relational Grammar 2. David Perlmutter and Carol Rosen, eds. Chicago: University of Chicago Press. 81-125.

Rosen, Carol G. 1984. The interface between semantic roles and initial grammatical relations. In Studies in Relational Grammar 2. David Perlmutter and Carol Rosen, eds. Chicago: University of Chicago Press. 38-77

Williamson, Janis. 1979. Patient marking in Lakhota and the Unaccusative Hypothesis. In Papers from the Fifteenth Regional Meeting of the Chicago Linguistic Society. Chicago: University of Chicago. 353-365.

Willson, Steve. 1985. Verb Agreement and Case Marking in Burushaski. MS.

KANASI: A Brief Grammar Sketch[1]

Ronald W. Pappenhagen

1 Introduction
2 Noun phrases
  2.1 Numerals and descriptive modifiers
  2.2 Genitive constructions
  2.3 Adpositions
3 Verbs
  3.1 Affixes
  3.2 Tense, aspect and mood
  3.3 Causation
4 Predicate nominals
5 Existential, locational and possessive sentences
  5.1 Existential verbs
  5.2 Existential sentences
  5.3 Locative sentences
  5.4 Possessive sentences
6 Verbal clauses
  6.1 Basic constituent orders
    6.1.1 Intransitive clauses
    6.1.2 Transitive clauses
    6.1.3 Indirect objects and obliques
    6.1.4 Quotation clauses
  6.2 Coding grammatical relations
    6.2.1 Word order
    6.2.2 Verb agreement
    6.2.3 Pronominal system
  6.3 Negation
  6.4 Questions
    6.4.1 Yes/No questions
    6.4.2 Information questions
  6.5 Comparison
7 Relative clauses
8 Conclusion
Notes
References

1 Introduction

Kanasi is a non-Austronesian language in the Indo-Pacific family of the Daga branch. It is spoken by approximately 2,000 people living in the Owen Stanley mountains in the Rabaraba subdistrict of the Milne Bay Province of Papua New Guinea. There is no previous linguistic literature related to Kanasi, though the Daga language has been described (Murane 1974) and work is presently being undertaken in Umanakaina, another language in the Daga branch (Evensen, forthcoming).

The data on which this sketch is based were collected under the auspices of the Summer Institute of Linguistics during nineteen months of residence in the village of Danobu between July 1982 and April 1985.

2 Noun phrases

2.1 Numerals and descriptive modifiers

In Kanasi, [2] numerals and descriptive modifiers follow their head nouns.

(1) Dágina éuda lo-tón.
skirt good take:off-3SG:PAST
He took off the good skirt.

(2) Né gagálo-na kámpaida wá-pa.
1SG talk-1SG:GEN short speak-1SG:FUT
I only have a little bit to say here.

(3) etá kérauda
garden:house small
small garden house

(4) bonó uára ía
pig many very
very many pigs

(5) tapároro natá ósi
week two only
only two weeks

(6) ónta áíanapa dubá
pot sooty:deposit black
black sooty deposit (on the bottom of a) pot

Kanasi does not use definite or indefinite determiners, though non-specific persons or things can be identified by the post-nominal particle dá:

(7) nené dá
bird some/another
some bird

The demonstrative **emá** follows its head:

(8) Né letá emá girúma-nten.
1SG letter DEMON write-1SG:PAST
I wrote this letter.

Note that in (8) if the order of noun and demonstrative is reversed a predicate nominal sentence results:

(9) Dimá? Emá letá.
what? DEMON letter
What is this? This is a letter.

Relative clauses also follow their heads:

(10) HEAD RESTRICTING SENTENCE
Letá [é wáúbe-'an] né wadá-na.
letter 2 send-2SG:PAST 1SG get-1SG:PAST
I got the letter (that) you sent.

The relative clause in Kanasi will be dealt with more fully in Sect.6.

This placement of numerals, descriptive modifiers, determiners, demonstratives and relative clauses after their head nouns is inconsistent with the canonical OV word order type (Sect. 5.1.2; Comrie 1981:89).

2.2 Genitive constructions

In Kanasi the possessor exclusively precedes the possessed item.

(11) Taii ebó-wa wanála
vine name-3SG:GEN vine:species
(The) vine's name is (i.e., is called) wanala

(12) Róni nuá-wa kadí-den.
Ron heart-3SG:GEN bad-3SG:PAST
Ron's heart was bad. (lit: Ron was upset.)

This placement of the genitive (possessor) in relation to the possessed item is consistent with general expectations for an OV language (Comrie 1981:89).

2.3 Adpositions

Kanasi is exclusively postpositional in that the semantic role of an oblique NP is specified by free standing

postpositions. The postpositions, however, are inflected by suffixation for the person and number of their arguments:

(13) iá osó-ve
tree on-3SG
in the tree (lit: on the tree)

(14) Tané bonó wá-wa osó-ve lapí-len.
later pig leg-3SG on-3SG trap-3SG:PAST
Later the pig's leg got caught in the trap.

(15) Dú bedé-ve wán-de.
house inside-3SG E:be/sit:AN-3SG:PRES
He is (sitting) in the house.

(16) ená dá bedé-ve
year another inside-3SG
next year

(17) Nó máin-imbo ví-sen.
1PL to-1PL come-3SG:PAST
He came to us.

This postpositional characteristic is consistent with the canonical OV word order type as discussed in Comrie (1981:89).

3 Verbs

3.1 Affixes

Kanasi employs suffixing on the verb to accomplish agreement with subject and object and also to signal the tense, aspect and mood. There are some marginal examples of prefixation used to code direct objects, but these are restricted to only one verb that we are aware of. Certainly, prefixing is not a salient feature of the language. The following are some straightforward examples of how Kanasi uses suffixing on the verb stem.

(18) Esá debáma nóse-n.
water big come:down-3SG:PAST
A lot of water came down. (i.e., had much pressure)

(19) Taii topúna téine-pi.
vine vine:species pull-3SG:FUT
He will pull the topuna vine.

(20) Né ná-'e-pa.
1SG eat-2:OBJ-1SG:FUT
I will eat you.

The only clear example of prefixing occurs with the verb 'give.' Here the prefixes always refer to recipients while the suffixes always refer to the agent, i.e., the giver.

(21) Káire emá e-né'e-pa.
sweet:potato DEMON 2SG-give-1SG:FUT
I will give this sweet potato to you.

(22) Káire emá wa-ná-pa.
sweet:potato DEMON 3SG-give-1SG:FUT
I will give this sweet potato to him.

In (21) and (22) the prefixes agree with the recipient object of the verb. The verb for 'give' is the only verb we have encountered to date that operates in this manner.

In addition to suffixation and the marginal case of prefixation outlined above, at least one Kanasi verb is suppletive depending on the number of the direct object. The verb meaning 'take' occurs in the form **wadé** if the patient is singular (ex. 23, 24) and in the form **otá** if the patient is plural (ex. 25):

(23) Manísi-wa wadé-n.
tobacco-3SG:GEN take:SG:item-3SG:PAST
He took his own tobacco.

(24) Vé wadé-'e-n.
sickness take:SG:item-2-3SG:PAST
You got sick. (lit: The sickness took you.)

(25) 'E natá otá-mo-n.
spear two take:PL:items-3PL:OBJ-3SG:PAST
He took two spears.

Most transitive verbs in Kanasi show no agreement with third person singular direct objects. Alternatively, one might say that the normal third person singular direct object marker is zero, as illustrated in (26) and (27):

(26) Iá tó-n.
tree chop-3SG:PAST
He chopped the tree (down).

(27) O ná-dina.
banana eat-1SG:PRES
I am eating the banana.

For some transitive verbs, however, third person singular direct objects are indicated with the suffix **-sa**:

(28) Bonó bó-sa-n.
pig die-3SG:OBJ-3SG:PAST
He killed the pig.

In sum, Kanasi uses only suffixes on the verb except in the case of 'give' which employs a prefix for the recipient object. This is consistent with the expectations for an OV language (Comrie 1981:90).

3.2 Tense, aspect and mood

The following is a preliminary analysis of the tense-aspect-mood system of Kanasi. A thorough-going study of the use of tense, aspect and mood in discourse will be necessary before any firm conclusions can be drawn concerning this complicated topic.

Kanasi grammaticalizes tense, aspect and mood through suffixes attached to the verb. The following are straightforward examples of past, present and future tenses. These examples are devoid of markers for aspect, mood, etc., as signified by the -0.

(29) Angelína imána náúkai wadé-n-0.
Angelina yesterday bowl take:one:item-3SG:PAST
Angelina took the bowl yesterday.

(30) Angelína náúkai táúma wadé-de-0.
Angelina bowl now take:one:item-3SG:PRES
Angelina is holding the bowl.

(31) Deáma á-lepi isíma wadé-pi-0
later go-3SG:FUT knife take:one:item-3SG:FUT
Later he will get the knife and bring it here.

vím-pi-0.
come-3SG:FUT

Aspect in Kanasi may be illustrated by the following common utterances that are used in leave-taking. When a person is about to go they say 'I am going.' Then the hearer usually responds with one of the following statements:

(32) Né wáine-dina.
1SG E:be/sit:AN-1SG:PRES1
I am sitting. (i.e., I will stay while you go.)

(33) Né wáine-pina.
1SG E:be/sit-1SG:PRES2
I am sitting.

Both examples are ostensibly present tense. But one would say (32) if one were standing as the other person were getting ready to leave and the speaker were **about to sit down**. On the other hand one would say (33) if one were **already** sitting down and would be planning to **continue to do so** while the other person left. So a better translation for (33) might be: 'I will continue to sit (here).'

3.3 Causation

Kanasi employs only lexical causatives rather than morphological or analytical causatives. Therefore the idea of causation is actually tied up in the verb itself without the addition of extra affixes or words.

(34)(Plane trip text.)
Nó asím-ba potá-len mái gagálo etá-pona.
1PL ear-1PL:GEN close:up-3SG:PAST NEG talk hear-NEG
Our ears were plugged (and) we could not hear.

Pléni asím-ba potá-sa-n.
plane ear-3PL:GEN close:up-3SG:OBJ-3SG:PAST
The plane plugged up our ears.

(35) Ambi dewá-len? Rón dewá-sa-n.
what do-3SG:PAST Ron do-3SG:OBJ-3SG:PAST
What happened? Ron did it.

The suffix -sa in (34) and (35) is merely a 3SG object marker and not a causative morpheme, as the following examples will show:

(36) Duádi-sa-pa.
bathe-3SG:OBJ-1SG:FUT
I will bathe him.

(37) Duádi-'e-pa.
bathe-2:OBJ-1SG:FUT
I will bathe you.

(38) *Duádi-sa-'e-pa.

While we conclude that Kanasi has lexical causatives, we acknowledge that other strategies could still be at work and that much more study needs to be done in this area.

4 Predicate nominals

Kanasi forms predicate nominal sentences by juxtaposing two noun phrases:

(39) Sá bodubódu waláma.
DEMON wet:season day
That is the wet season.

(40) Iá emá sabóma.
tree DEMON tree:species
This is a saboma tree.

(41) Aúpa mám-ba Omoasi.
Aupa father-3SG:GEN Omoasi
Aupa's father is Omoasi.

(42) Rábawai baráú apána.
Rabawai sorcery man
Rabawai is a sorcerer.

(43) Apána sá baráú apána.
man DEMON sorcery man
That man (just referred to) is a sorcerer.

5 Existential, locational and possessive sentences

5.1 Existential verbs

In Kanasi the existential, locational and possessive sentences are formed by combining a noun phrase and a locative phrase (in either order) followed by a verb. We call this verb an 'existential verb.' Existential verbs are not technically copulas, since they do not occur in

predicate nominal constructions. The choice of existential verb in any construction is determined by the class of the noun phrase of which it predicates existence. The following is a list of existential verbs and the noun classes to which they belong.

Verb Stem	**Gloss**	**Noun Phrase Class**
wán-	E:sit	animate
viné-	E:be	inanimate
ené-	E:stand	animate and inanimate and tall
seú-	E:hang	inanimate and hanging

5.2 Existential sentences

In Kanasi the existential sentence is formed with an indefinite noun phrase normally followed by a locative phrase and then the existential verb:

NP (LOC) E:VERB
[indef]

(44) Mápo nené dá ebó-wa bináma
long:ago bird another name-3SG:GEN bird:hornbill
Once upon a time there was a hornbill bird

só-wa kóko natá wám-oi-ena.
friend-3SG:GEN bird:cassowary CONJ E:sit:AN-3PL-PAST
and his friend a cassowary bird.

(45) Boní emá wáím-oi bá día?
ghost here E:sit:AN-3PL:PRES or NEG
Are there ghosts here or not?

Boní uára wáím-oi!
ghost many E:sit:AN-3PL:PRES
There are many ghosts!

(46) Boní uára irá'i-na bedé-ve wáím-oi.
ghost many garden-1SG:GEN in-3SG E:sit:AN-3PL:PRES
There are many ghosts in my garden.

(47) Sopí viné-de.
soap E:be:INAN-3SG:PRES
There is soap.

(48) Wáita seú-moi.
bush:mango E:hang-3PL:PRES
There are bush mangoes.

5.3 Locative sentences

The locative sentence is formed with a definite or an indefinite noun phrase followed by a locative phrase and then the existential verb:

NP (LOC) E:VERB
[def/indef]

(49) Eló emá ká ató-ve wán-de.
dog DEMON table under-3SG E:sit:AN-3SG:PRES
This dog is under the table.

(50) Buká sá waká osó-ve viné-de.
book DEMON table on-3SG E:be:INAN-3SG:PRES
The book is on the table.

(51) Ama-na dú-na bedé-ve viné-de.
hammer-1SG:GEN house-1SG:GEN in-3SG E:be:INAN-3SG:PRES
My hammer is inside my house.

5.4 Possessive sentences

In Kanasi the most common possessive sentence type is parallel to the existential sentences described above, except that the nominal whose existence is predicated is possessed:

(52) Mé obá'am-pa uára wáim-oi.
3SG children-3SG:GEN many E:sit:AN-3PL:PRES
He has many children.

(53) Né dú-na ené-de.
1SG house-1SG:GEN E:stand-3SG:PRES
I have a house.

Example (52) could be paraphrased 'His many children exist', and (53) might be paraphrased 'My house exists'. There are also some less common ways to frame possessive sentences, as illustrated in (54) through (56):

(54) Nó máin-imbo iá dia.
1PL to-1PL wood NEG
We do not have wood.

(55) Iá dia mám-pe.
wood NEG to-3SG
He does not have any wood.

(56) Ama ínta ná? Mamasia mám-pe
hammer way where Mamasia to-3SG
Where is the hammer? Mamasia has it.

(vine-de).
(E:be:INAN-3SG:PRES)

These forms do not employ an existential verb. As with many languages (Clark 1978), these possessive sentences are formally parallel to locational sentences, as described in section 4.3., with the possessor coded as a locative NP (**mámpe** 'to him', **mámpu** 'to them,' etc.). Perhaps it is significant that in all the contexts for (54), (55) and (56) the 'owner' is not really owning but only has something temporarily.

The last example of a possessive sentence shows yet another form:

(57) Mé mám-ba dia, (mái wám-pona, dia).
3SG father-3SG:GEN NEG (NEG E:sit:AN-NEG NEG)
He has no father, (He does not exist/is not alive.).

This example illustrates that in the negative, the existential verb
may be omitted.

6 Verbal clauses

6.1 Basic constituent orders

6.1.1 Intransitive clauses. The basic constituent order in intransitive clauses is SV:

(58) Róni né natá á-nten.
Ron 1SG CONJ go-1:PAST
Ron and I went.

(59) Sikóko púnde-n?
water:spout break-3SG:PAST
Did the water spout break?

(60) Bónim-o víne-na tadín-o itú-nten.
Bonima-LOC come-1SG:PAST night-LOC sleep-1:PAST
I arrived in Bonima at night and slept (there).

6.1.2 Transitive clauses. The basic constituent order in transitive clauses is SOV:

(61) Stíven áia badówa wá-de.
Stephen language strong speak-3SG:PRES
Stephen is speaking well.

(62) Né letá emá girúma-nten.
1SG letter DEMON write-1SG:PAST
I wrote this letter.

(63) Boní omán áúbe-n uá-wa epé-n.
ghost rock throw-3SG:PAST head-3SG:GEN hit-3SG:PAST
The ghost threw a rock (and) hit her in the head.

Kanasi can be characterized as a very rigid SOV language, since variant orders only occur under pragmatically very marked conditions, such as false starts and restatements. I will not argue further here for SOV as the basic constituent order in Kanasi clauses.

6.1.3 Indirect objects and obliques. Indirect objects and other obliques appear in any position relative to the verb:

(64) Riá-sa-'a né máín-awo vím-pi.
slide-3SG:OBJ-2SG:IMP 1SG to-1SG come-3SG:FUT
Slide it towards me.

(65) Ináko mám-pe á-íta.
mommy to-3SG go-2SG:IMP
Go to (your) mommy.

(66) Róni né natá á-nten sú esá mám-pe.
Ron 1SG CONJ go-1:PAST up water to-3SG
Ron and I went up to the water (tap).

(67) Esá mám-pe víne-n Róni dewá-sa-n.
water to-3SG come-1:PAST Ron do-3SG:OBJ-3SG:PAST
When we came to the water (tap), Ron did it.

6.1.4 Quotation clauses. With respect to quotations, the quoted information follows the verb 'say' so that one could state that with these sentences the pattern is not SOV but rather SVO:

(68) Tané binámа wá-n-ase
CONJ bird:hornbill say-3SG:PAST-QUOTE
Then the hornbill bird said:

Máí é báúta duádi-ta'a.
NEG 2 first bathe-2:SG:IMP
'No, you bathe first.'

6.2 Coding grammatical relations

Kanasi uses a variety of means to distinguish the semantico-syntactic roles of **S** (the single argument of an intransitive clause), **A** (the most agentive argument of a transitive clause) and **P** (the most patient-like argument of a transitive clause).

6.2.1 Word order. The main method Kanasi employs to distinguish the grammatical relations is word order. In the transitive clause the role order is always A-P-Verb:

(69) Alan á-len.
Alan go-3SG:PAST
Alan left.

(70) Alan Báti epé-n.
Alan Bati hit-3SG:PAST
Alan hit Bati.

(71) Alan káire ná-n.
Alan sweet:potato eat-3SG:PAST
Alan ate the sweet potato.

(72) *Káire Alan ná-n.

6.2.2 Verb agreement. The roles A and P may be distinguished through morphological verb agreement. The verbal constituent order is:

STEM-OBJECT:MARKER-SUBJECT:MARKER

(73) Bonó adé-'e-pi!
pig bite-2:OBJ-3SG:FUT
The pig will bite you!

(74) Alan posí epá-mo-de.
Alan ball hit-3PL:OBJ-3SG:PRES
Alan is kicking the balls.

(75) Sedá ewá-mo-na.
young:people see-3PL:OBJ-1SG:PAST
I saw the young people.

6.2.3 Pronominal system. The roles S and A can be coded with pronouns and with full noun phrases whereas P cannot. The P can only be coded as a full noun phrase or as a verbal suffix in a pragmatically unmarked clause. It is also important to note that the pronouns can only be used with human S's and A's. If the construction has a non-human S or A then a noun phrase will be used in place of the pronoun.

SUBJECT PRONOUNS AND OBJECT VERB AGREEMENT MARKERS IN KANASI

	S & A		P	
	Subject Pronouns		Object Verb Agreement	
	Singular	Plural	Singular	Plural
1st person	né	nó	-na	-no
2nd person	é	é	-'e	unknown
3rd person	mé	mó	0,-sa	-mo-

(76) Mé adá-na-en.
3SG:HUM bite-1SG:OBJ-3SG:PAST
He bit me.

(77) Nó ináwa'a máí keró wadá-no-pona.
1PL:HUM healthy NEG sickness get-3PL:OBJ-NEG
We are not sick. (lit: The sickness has not got us.)

(78) Né ná-'e-pa.
1SG:HUM eat-2:OBJ-1SG:FUT
I will eat you.

(79) Otó-pa duádi-sa-de.
child-3SG:GEN bathe-3SG:OBJ-3SG:PRES
She is bathing her child.

(80) Náúkai mité-de.
bowl wash-3SG:PRES
She is washing the bowl.

(81) Náúkai mitá-mo-de.
bowl wash-3PL:OBJ-1SG:FUT
She is washing the bowls.

Example (79) illustrates the use of the -sa as the 3SG object marker. But in (80) there is nothing to mark the 3SG object. This presents an unresolved problem about determining which verbs will have an object marker for 3SG and which will not.

The following chart displays the verbs that use -sa and those that do not use it. We have divided the Kanasi verbs into two classes based on the 3SG past tense endings.

Class II that never take -sa-	Class II that do take -sa-	Class II with -sa- included have Class I endings	All Class I never take -sa-
á-len 'he left'	lagá-len 'he lied'	lagá-sa-n 'deceived him'	evé-n 'he saw it'
nú-len 'he cooked it'	potá-len 'it closed up'	potá-sa-n 'closed it up'	wadé-n 'he took it'
itú-len 'he slept'	dewá-len 'it happened'	dewá-sa-n 'he did it'	wané-n 'he gave it'
rovó-len 'he jumped'	duádi-len 'he bathed'	duádi-sa-n 'bathed him'	dié-n 'asked him'
iní-len 'it grew'	kadí-den 'it broke'	kadí-sa-n 'he broke it'	é-n 'he left it'
nó-sen 'it fell'	bó-len 'he died'	bó-sa-n 'he killed it'	wá-n 'he said it'
ví-sen 'he came'	día-len 'it ended'	día-sa-n 'finished it'	ná-n 'he ate it'
	dabúa-len 'he is bored'	dabúa-sa-n 'he bored him'	ínse-n 'perceived it'
	riá-len 'it slid'	riá-sa-n 'he slid it'	rovíne-n 'he turned it'
	form unknown	sáú-sa-n 'helped him'	téine-n 'he pulled it'
	form unknown	wadákea-sa-n 'clawed him'	áí-n 'put it inside'
			'ené-n 'speared it'
			tó-n 'he cut it'
			itú-n 'put it down'

		ié-n 'he came down'
		odé-n 'he came up'
		ené-n 'he stood up'
		wáí-n 'he sat'

The pattern that emerges is that Class I verbs are transitive and Class II verbs are intransitive. This hypothesis is strengthened by the fact that the addition of **-sa** (3SG object marker) to most Class II (intransitive) verbs makes those verbs have Class I (transitive) endings.

6.3 Negation

The negation strategy for Kanasi is a negative particle and/or suffix. There is no one way to form negatives in Kanasi, but the most common method is the following:

máí + VERB-pona = negative in past tense

máí + VERB = negative in future tense

form unknown = negative in present tense

When the **-pona** suffix is employed, the verb, depending on its class, can be affected. A Class I verb (i.e., with 3SG:PAST suffix **-n**) will lose its person marker when the suffix **-pona** is attached. A Class II verb (i.e., with 3SG:PAST suffixes **-len** or **-sen**) will retain the person marking. For these Class II verbs the tense, of course, is retained since the **-pona** automatically means 'past tense.'

It is also worth noting that when the person marker is dropped with the Class I verbs, it is **not** transferred to the negative element. It is just that the marking is gone altogether so that the forms all look alike.

The following are examples of these common negatives illustrating the two verb classes:

	Class I Verb	**Class II Verb**
PAST TENSE		
(82)	Né wá-na. 1SG say-1SG:past I said it.	(83) Né itú-nte-n. 1SG sleep-1SG-PAST I slept.
(84)	Né máí wá-pona. 1SG NEG way-NEG I did not say it.	(85) Né máí itú-nta-pona. 1SG NEG sleep-1SG-NEG I did not sleep.
FUTURE		
(86)	Né wá-pa. 1SG say-1SG:FUT I will say it.	(87) Né itú-ntea-pa. 1SG sleep-1SG-FUT I will sleep.
(88)	Né máí wá-pa. 1SG NEG say-1SG:FUT I will not say it.	(89) Né máí itú-ntea-pa. 1SG NEG sleep-1SG-FUT I will not sleep.
PRESENT		
	forms unknown	

The negative particle **máí** is used alone (i.e., without **-pona**) in the predicate nominal construction.

(90) **Né baráú apána.**
1SG sorcery man
I am a sorcerer.

(91) **Né máí baráú apána.**
1SG NEG sorcery man
I am not a sorcerer.

The suffix **-pona** is also employed for the irrealis mode, especially in contrary-to-fact constructions:

(92) Aúpa óme-n.
Aupa fall:AN-3SG:PAST
Aupa fell.

(93) Aúpa **máí** óme-pona.
Aupa NEG fall:AN-NEG
Aupa did not fall.

(94) Aúpa óme-pona.
Aupa fall:AN-COND
Aupa almost fell (but did not).

Another method for forming negatives is the particle **si'a**. At present we do not see any difference between the meanings for the **si'a** and the **máí + -pona** forms.

Class I	**Class II**
(95) Wá sí'a. say NEG It was not said. (Nobody said it.)	(96) Itú sí'a. sleep NEG Nobody slept.
(97) E wá sí'a. 2 say NEG You did not say it.	(98) E duádi sí'a. 2 bathe NEG You have not bathed.

Another method of forming the negative, which may be related to the **si'a** particle, is the suffix **-si**:

Class I	**Class II**
(99) Né wá-si-a-nten. 1SG say-NEG-go-1:PAST I did not say it.	(100) Né itú-si-a-nten. 1SG sleep-NEG-go-1:PAST I did not sleep.

This is not a form that we have often encountered and for the most part has come to our attention only through direct elicitation. Future research will reveal the extent to which **-si** is a productive negative strategy, and whether any nuances of meaning are associated with it that are not revealed in the elicited translations.

We have examined the negative particles and suffixes **máí, -pona** and **-si, si'a**. Now we shall turn our attention to the negative particle **día**. This particle appears most commonly in questions and in answers to questions. When it is used to answer a question, it occurs in conjunction with the constructions discussed above:

(101) E ái-ten bá día? Día, máí á-nte-pona.
2 go-2SG:PAST or NEG NEG, NEG go-1SG-NEG
Did you go or not? No, I did not go.

(102) Mé á-len bá día? Máí á-le-pona, día.
3SG go- SG:PAST or NEG NEG go-3SG-NEG, NEG
Did he go or not? He did not go.

Example (102) illustrates that día can come at the end of the answer as well as at the beginning as in (101) with no apparent effect on the meaning.

The word **día** also appears as a prominent part of some negative possessive sentences.

(103) Nó máín-imbo iá día.
1PL to-2PL wood NEG
We do not have any wood.

(104) Iá día mám-pe.
wood NEG to-3SG
He does not have any wood.

(105) Mé mám-ba día.
3SG father-3SG:GEN NEG
He has no father.

In sum, the unmarked negation strategy in Kanasi is the particle **máí**, often occurring in conjunction with negative suffixation of one sort or another. Also the finite verb can lose the person marker when the **-pona** negative suffix is attached, and this is conditioned by the class of the verb in question. But the lost person marker does not attach itself to the negative particle. Since Kanasi predominantly employs a negative particle strategy for negation, it is consistent with the expectation for an SOV language.

6.4 Questions

6.4.1 Yes/No questions. There is no obvious intonational difference between yes/no questions and unmarked declarative sentences:

(106) Sikóko púnde-n?
water:spout break-3SG:PAST
Did the water spout break?

Further study may reveal some subtle intonational characteristic of yes/no questions (e.g. rising intonation elsewhere other than sentence-finally).

The major yes/no question strategy involves the use of a sentence-final tag particle:

(107) E ái-ten bá (dia)?
2 go-2SG:PAST or (NEG)
Did you go or not?

(108) Ná-'(a) ó?
eat-2SG:IMP QTAG
Do you want to eat it?

Like declarative sentences, this yes/no question structure has sentence-final falling intonation. The tag particle **ó** identifies this sentence as a question.

The use of a sentence final question particle is consistent with the OV language type as described by Greenberg (1966:10).

6.4.2 Information questions. In information questions the question word occurs wherever the corresponding non-question word would occur in a corresponding declarative sentence.

Examples of sentence-initial question words:

(109) Dimá wadé-n?
What? take:one:item-3SG:PAST
What did he take?

(110) Midá rába itú-n?
Who? rubber:pipe put:one:item-3SG:PAST
Who put the pipe (in the water)?

(111) Ná á-len?
Where? go-3SG:PAST
Where did he go?

(112) Ambi dewá-len?
How? do-3SG:PAST
How did it happen?/What happened?

Examples of sentence medial question words:

(113) Tané nó dimá wadá-te'e odá-ntata?
CONJ 1PL what? get-1PL:ASP feast-1PL:FUT
But what will we be able to get for the feast?

(114) E obán-'a ámbi wáim-oi?
2 children-2SG:GEN how:many? E:sit:AN-3PL:PRES
How many children do you have?

(115) E ná áí-ti'ana?
2 where? go-2SG:PRES
Where are you going?

(116) Waláma dáinea' é áí-ti'ana?
day which? 2 go-2SG:PRES
Which day are you going? (i.e., when...?)

The following are examples of question words in sentence final position. It is interesting to note that the question word appears sentence finally only in sentences which have no verb overtly present.

(117) Ama ínta ná?
hammer road where?
Where is the hammer?

(118) Mé ebó-wa ámbi?
3SG:HUM name-3SG:GEN how?
What is his/her name?
(lit.: How is his name?)

(119) Mé ebó-wa midá?
3SG:HUM name-3SG:GEN who?
What is his name?
(lit.: Who is his name?)

(120) Emá dimá?
DEMON what?
What is this here?

(121) Omó dimá?
sun what
What time is it?

(122) Nitú dáinea'?
thing which?
What is (that) thing?

(123) E apá(na) dáinea'?
2 man which?
Who are you?

The data show that the information question words in Kanasi can appear sentence initial or elsewhere, as would be expected in an OV language (Greenberg 1966:111).

6.5 Comparison

Kanasi has no grammaticalized way of forming comparatives. The following example illustrates how comparison is accomplished paraphrastically:

(124) Emá debáma emá kérauda.
DEMON big DEMON small
This is big (and) this is small.

Since Kanasi does not have a comparison strategy employing a standard of comparison, a marker and an adjective of comparison, nothing can be said about how it meets the expectations for an OV language with respect to these elements (Comrie 1981:92).

7 Relative clauses

Relative clauses are not very common in Kanasi. The following are the only two examples of relative clauses we have encountered in the corpus used for this sketch. In the following examples, the head is underlined and the relative clause is bracketed:

(125) **Letá** [é wáúbei-'ana] né wadá-na.
letter 2 send-2SG:PAST 1SG get-1SG:PAST
I got the letter (that) you sent.

(126) Rón Dién Stíven **otó** **kérauda** [ebó-wa
Ron Jane Stephen child small name-3SG:GEN
Ron, Jane, Stephen and the little baby, whose

ámbi ne móa-na] dó udú'udu níapo.
how 1SG not:know-1SG also all greeting
name I do not know, greetings to all of you.

In both of these examples the restricting clause follows the head. The strategy in (125) is the gap type in that the relativized element has been omitted from the restricting clause and there is no relative pronoun. The head and restricting clause are merely juxtaposed. It also should be noted that in (125) the object is being relativized.

The example in (126) is a bit unusual in that it contains no matrix verb. Still, it can be seen that the restricting clause follows the head as in (125). Example (126) relativizes the possessive and instead of using the gap strategy (126) uses the pronoun retention strategy. As in (125), there is no relative pronoun but the word **ámbi** has been included. Strictly speaking **ámbi** is an interrogative word.

A second hypothesis for example (126) is that it may not even be a relative clause. Rather it may be only a parenthetical expression, an aside, so to speak. If this be so, the ámbi may indeed be an interrogative and so the speaker was in effect asking what the child's name was.

8 Conclusion

This sketch has shown that in most respects Kanasi is consistent with the expectations for the SOV language type, as outlined in Greenberg (1966). Kanasi is inconsistent with the SOV norm only in its post-nominal positioning of the relative clause. However, relative clauses are quite rare in Kanasi discourse, and further research may reveal patterns other than those observed to date.

Notes

1. The basic work on this paper was completed to satisfy in part the course requirements for Linguistics 505 (Typology and Discourse) taken during the summer session 1985 of the Summer Institute of Linguistics at the University of North Dakota in Grand Forks. I would like to express appreciation to my colleagues Dr. Tom Payne, Dr. David Tuggy, Mr. Steve Knapp and Mr. Rick Nivens, all of S.I.L., for their guidance in producing this brief grammar sketch. Appreciation is also extended to the many Kanasi friends who patiently recorded and helped transcribe and gloss Kanasi text material essential to this analysis.

2. Abbreviations and orthographic conventions used throughout this sketch are the following:

*	=	grammatically unacceptable form
:	=	separates two English words glossing one Kanasi form
( )	=	inclusion optional
1SG	=	first person singular
1PL	=	first person plural
2	=	second person singular or plural
2SG	=	second person singular
2PL	=	second person plural
3SG	=	third person singular
3PL	=	third person plural
A	=	item highest in agentivity in a transitive clause
AN	=	animate
ASP	=	aspect
COND	=	conditional
CONJ	=	conjunction
def	=	definite
DEMON	=	demonstrative
E	=	existential verb
FUT	=	future tense
GEN	=	genitive
HUM	=	human
IMP	=	imperative
INAN	=	inanimate
indef	=	indefinite
lit.	=	literally
LOC	=	locative
NP	=	noun phrase
NPrel	=	relativized noun phrase
OBJ	=	object
OBL	=	oblique object
P	=	item most patient-like in a transitive clause
PAST	=	past tense

PL = plural
PRES = present tense
QTAG = question tag particle
S = single argument in an intransitive clause
SG = singular

The symbols used to represent Kanasi forms in this paper generally follow the IPA conventions, with the following exceptions:

l = voiced alveolar lateral continuant.

r = flap r [r]

' = glottal stop

á = stress

Dashes are used in Kanasi forms cited within the text to indicate whether each form is a prefix, suffix or root. For example **e-** indicates a prefix, **-sa** indicates a suffix, and **wadé** (no dashes) indicates a root. This convention obscures the fact that verb roots and many suffixes never occur without some suffix following them. So, for example, **wadé** would never occur as a free form in natural discourse, and a well-formed word would never end with the suffix **-sa**. Nevertheless, in keeping with common practice, we have opted for this convention, as any alternative seems even more opaque.

Most of the examples appearing in this paper are drawn from an extensive body of texts collected by myself and my wife, Jane Pappenhagen, although common expressions are used when needed to illustrate completely straightforward forms.

References

Clark, Eve C. 1978. Locationals: existential, locative and possessive constructions. Universals of human language, ed. by Joseph H. Greenberg, 85-126. Stanford: Stanford University Press.

Comrie, Bernard. 1981. Language universals and linguistic typology. Chicago: University of Chicago Press.

Evensen, Sigmund. (Forthcoming) Umanakaina grammar sketch. Ukarumpa, Papua New Guinea: Summer Institute of Linguistics.

Greenberg, Joseph. 1966. Some universals of grammar with particular reference to the order of meaningful elements. Universals of language, ed. by J. Greenberg. Cambridge, Mass: M.I.T. Press. pp. 73-113.

Murane, Elizabeth. 1974. Daga grammar: morpheme to discourse. Dallas: Summer Institute of Linguistics.

THE LANGUAGES OF THE ZAMBALES MOUNTAINS: A Philippine Lexicostatistic Study

John Wimbish

1 Overview
2 Introduction
3 Data gathering and analysis
4 The Zambales Ayta Language Tree
5 Summary
Appendix
Notes
References

1 Overview

A lexicostatistic survey was conducted in early 1985 as part of a comprehensive study of the languages spoken in the Zambales Mountains of the Philippines. The survey involved the collection and analysis of some forty word lists, resulting in the definition of six distinct languages in an area whose linguistic situation had previously been relatively unknown. This paper presents these survey findings in the form of a language tree, and introduces a general-use computer program which greatly enhanced the accuracy and speed of the lexicostatistic analysis.

2 Introduction

A yet unresolved mystery in anthropological studies has been the history of the Philippine Negrito community, a people of dark complexion, short stature, and black, curly hair. Except for their preference of dwelling in the more mountainous areas of the country, the only trait separating these people from other Filipinos is their physical makeup. No cultural bond exists between them, nor are there any linguistic commonalities. Rather, their languages and customs tend to be related more to those of their geographic neighbors. All attempts to find such characteristics to unite the different Negrito groups and to distinguish them from the lowland communities have ended in frustration (Headland 1975:49-50).

The Zambales mountains (Figure 1) are home for some ten thousand Negritos. These Zambales Negritos

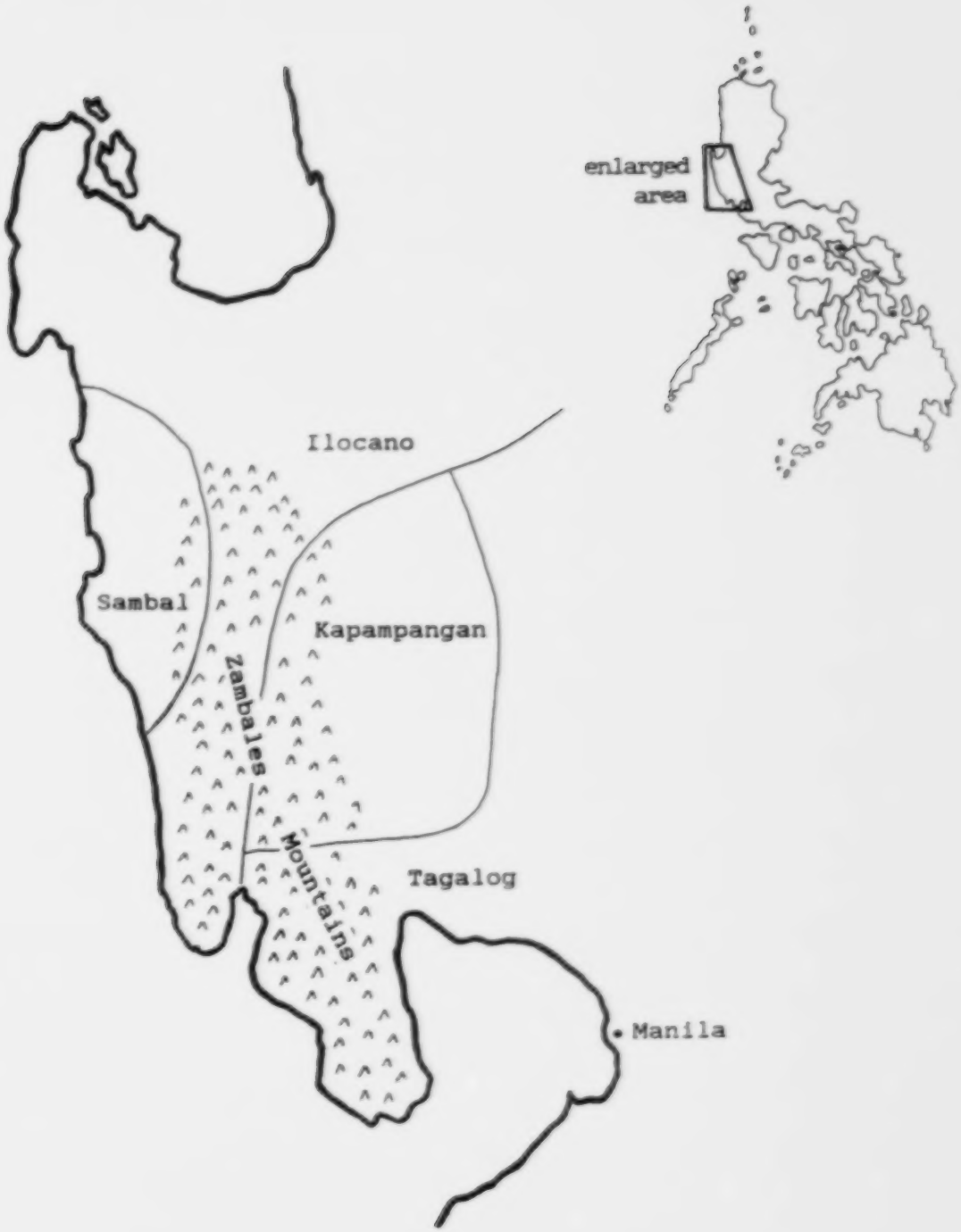


Figure 1 - Location of Zambales Mountains
and Trade Languages

(alternately, Aytas) were once suspected to all be speakers of a language closely related to Sambal, a nearby lowland language (Busenitz 1973). Later studies (Kyle 1977, Savage 1982) revealed the existence of several Ayta dialects, but due to limits on the number of locations they were able to visit, the geographic boundaries of these dialects remained undefined. The need for a comprehensive survey was established when attempts to use materials developed in the Sambal language for literacy instruction met with comprehension difficulties among the Aytas; the teachers being informed that there were nearly as many Ayta dialects as villages, and that few Ayta could understand Sambal.

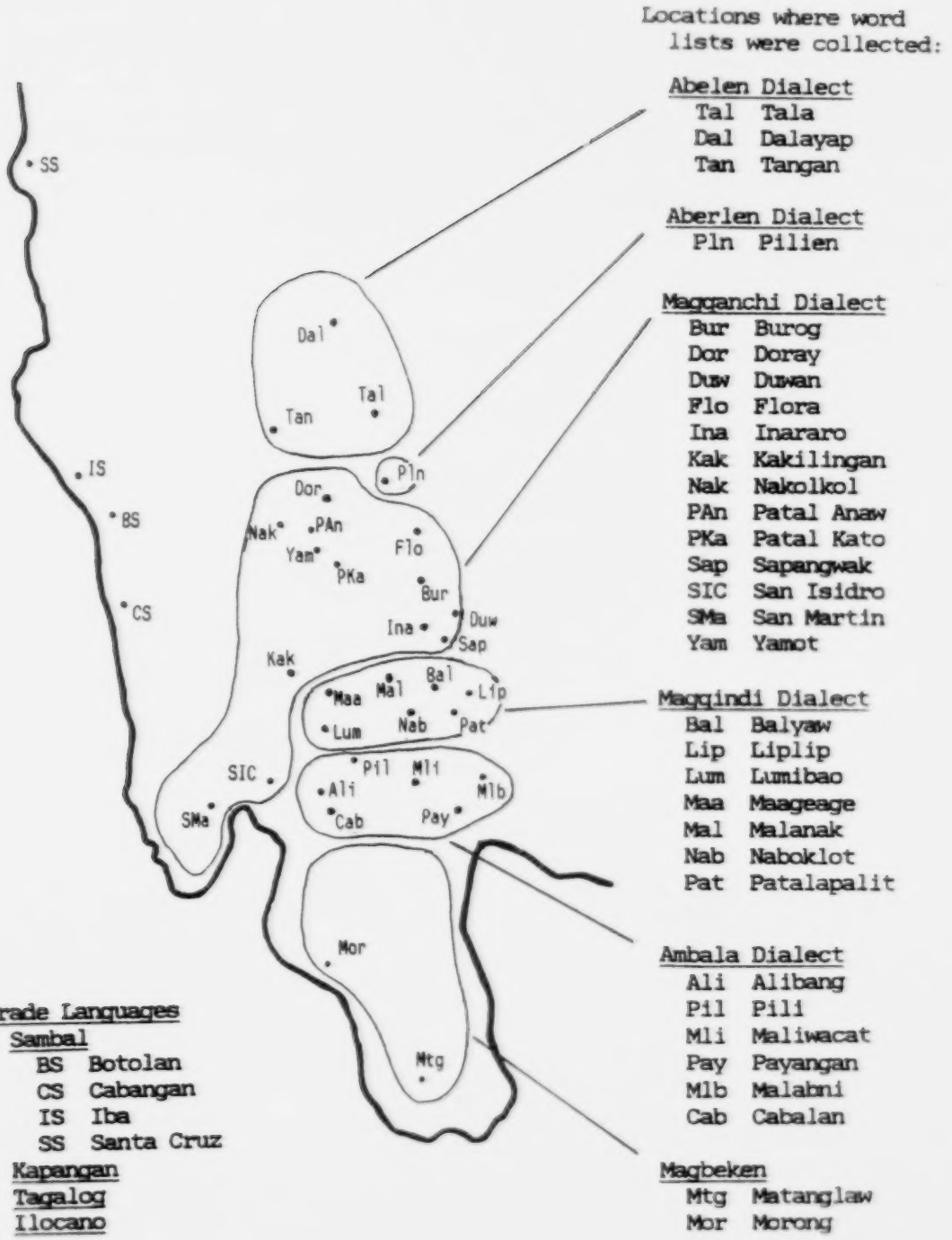
With the goal of developing adequate language materials for all of the people of the Zambales mountains, the Summer Institute of Linguistics (SIL) planned a two-part survey. The first phase would be a lexicostatistical study, entailing the gathering and comparison of word lists from every location possible, in order to gain a picture of the overall language situation. The second phase would then consist of Casad-type text testing and sociolinguistic study, to further define each language's geographic boundaries and to indicate the Ayta's understanding and use of the trade languages in each area. What follows is a presentation of the results of the initial phase.

3 Data gathering and analysis

The collection of word lists has often been the first step in linguistic survey procedure, providing a general and rough classification of the dialects in the target area. Experience among SIL members in Philippine comparisons has generated two empirical definitions for the interpretation of cognate scores between lists[1]. The first states that two lists represent the exact same dialect if their vocabularies are greater than 95 percent cognate. The second claims that two lists represent mutually unintelligible languages[2] if their vocabularies are less than 60 percent cognate. Lists with percentages between the two figures can only be tentatively classified, relying on text testing to give a more accurate assessment of the intelligibility between the two locations.

The Reid Word List, consisting of 372 items tailored for use in the Philippines, was collected in 40 locations[3], as represented in Figure 2. The comparison of these lists created a challenge: the mathematics predicted nearly eight

Figure 2 - Ayta Dialect Boundaries



hundred comparisons, requiring over six months of man-time to perform. As an alternative, the computer program WORDSURV[4] was developed.

WORDSURV allows the comparison of up to ninety word lists. Each list may contain as many as 400 items. A list is entered once through the program's editor, during which time the analyst groups the elicited forms into cognate sets with forms from other lists he has previously entered. Once a list has been entered, its cognate percentages with the other lists on file may be computed. This computation requires only a few minutes of computer time.

The analyst then arranges the cognate percentages into matrix form, not unlike distance charts on road maps. In the sample matrix in Table 1, the dialect spoken in the village of Doray shows 93 percent of its basic vocabulary to be cognate with the dialect spoken in Burog. Dialect names are written in two sections, showing language grouping and location: thus Magqanchi-Doray refers to the variety of the Magqanchi language spoken at Doray.

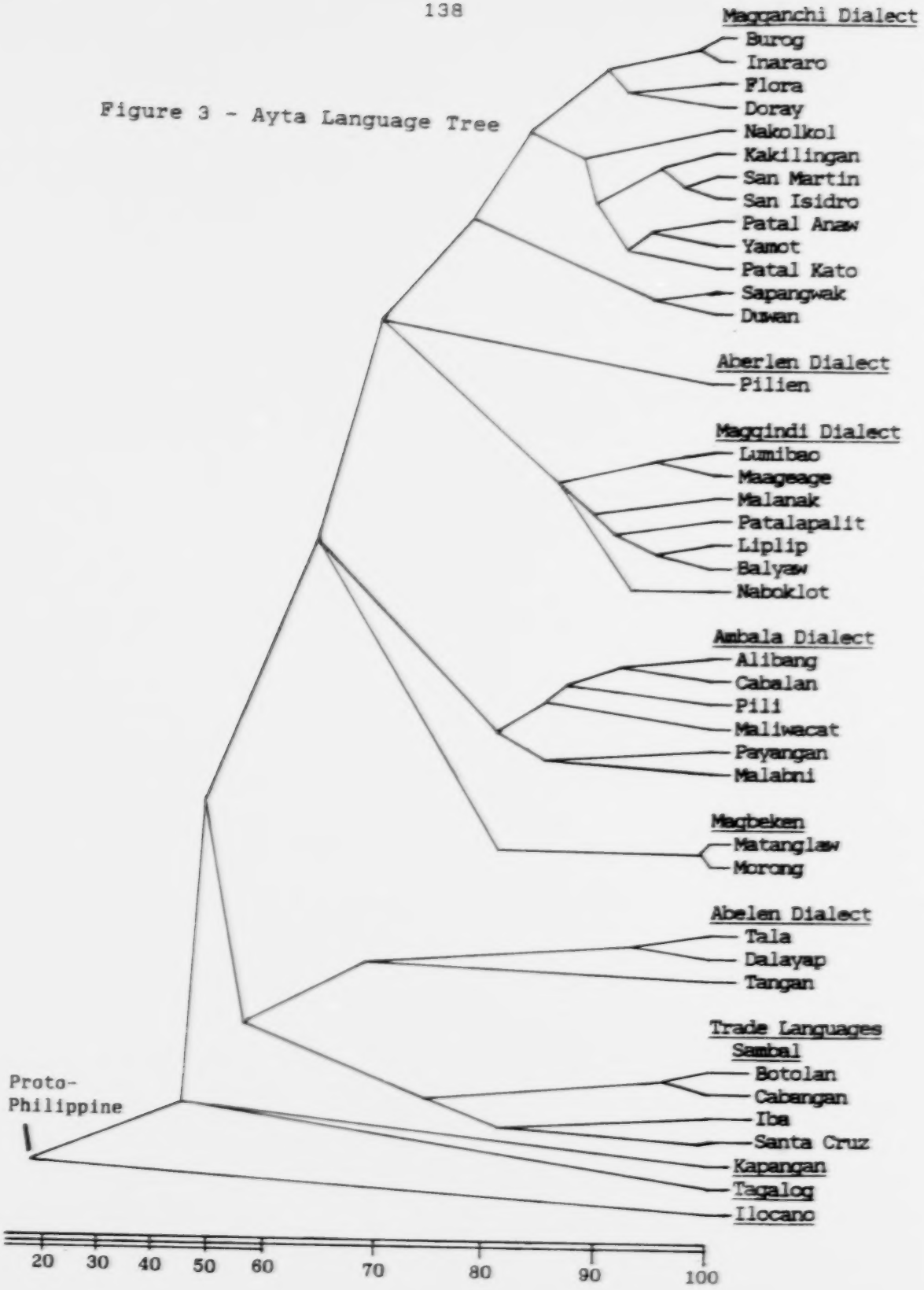
Magqanchi-Doray
92 Magqanchi-Flora
93 93 Magqanchi-Burog
90 90 98 Magqanchi-Inaararo
89 86 88 88 Magqanchi-Nakolkol

Table 1 - A Sample Matrix

Several methods of deriving a language tree from a cognate matrix are discussed by Walton (1977), who found that the "lowest-number" method generated the most accurate results in his study of 122 Philippine languages. This method, combined with the author's knowledge of geographic and lexical borrowing patterns in the area, was employed in the analysis.

It is recognized that a derivation based on shared cognate percentages lacks much of the accuracy of the tree that would be produced through a rigorous application of the comparative method. But it is also recognized that such an application of the comparative method with these word lists would require an extensive time investment. The tree presented in this paper is intended to merely provide a rough picture of the development of these Ayta languages, with the understanding that further study would undoubtedly generate refinements.

Figure 3 - Ayta Language Tree



4 The Zambales Ayta Language Tree

The circled areas of Figure 2 indicate the proposed major groupings derived from the word list comparisons. The six languages defined are:

Ayta Language	Population[5]
1. Mag-anchi	4166
2. Aberlen	184
3. Mag-indi	2483
4. Ambala	1657
5. Mabeken	572
6. Abelen	2000

TABLE 2 - Major Ayta Language Groupings

It is anticipated that dialects from villages within a given language grouping will exhibit a fair amount of intelligibility; village dialects separated by these language boundaries should prove inherently unintelligible.

The tree derived from the cognate matrix is presented in Figure 3. The cognate percentages are listed in the appendix.

5 Summary

The Zambales Mountains contain six different Negrito languages, Abelen, Magqanchi, Aberlen, Magqindi, Ambala, and Mabeken. An assessment of the ability of the speakers of these languages to understand their neighbors and the trade languages is the goal of the second phase of the survey. Though one cannot predict intelligibility on the basis of cognate scores alone, the groupings presented should provide a starting point for further examination of language use patterns among these Zambales Ayta.

APPENDIX

AYTA LEXICOSTATISTIC MATRIX

Accuracy codes for each word list are given in brackets; see note 3 for an explanation of their meaning. The cognate percentage between two lists is read as the intersection between their names. For example, the Abelen-Dalayap and Magqanchi-Doray lists are 63 percent cognate.

[C] Sambal-Iba
[C] 81 Sambal-Santa Cruz
[B] 75 74 Sambal-Cabanggan
[A] 76 74 96 Sambal-Botolan
[B] 44 44 45 46 Ilocano
[B] 41 44 44 44 47 Pangasinan
[B] 48 47 54 53 45 38 Tagalog
[C] 44 43 48 47 35 34 53 Pampanga
[C] 57 56 64 66 38 38 44 37 Abelen-Dalayap
[B] 59 57 86 68 38 40 44 39 93 Abelen-Tala
[C] 60 59 68 73 39 40 45 42 69 78 Abelen-Tangan
[B] 60 58 73 76 36 37 50 48 59 61 65 Magqanchi-Flora
[A] 60 59 75 78 40 37 50 49 63 62 67 93 Magqanchi-Buraq
[A] 60 59 75 78 40 38 49 47 63 63 67 90 98 Magqanchi-Inasaro
[B] 66 64 79 82 42 39 51 51 63 63 69 92 93 90 Magqanchi-Doray
[B] 65 64 82 86 40 39 51 47 65 66 72 86 88 88 89 Magqanchi-Kakalkal
[B] 67 80 84 88 41 42 51 46 66 68 72 85 89 89 87 89 Maganchi-Kakilingan
[A] 64 64 83 85 40 40 50 47 65 67 72 83 86 87 87 88 95 Magqanchi-San Martin
[A] 65 64 83 87 41 41 51 47 65 67 72 85 87 87 89 90 95 97 Magqanchi-San Isidro
[A] 58 58 69 73 36 36 48 50 80 61 64 88 91 88 87 84 83 81 83 Magqanchi-Sapangpak
[B] 55 63 55 69 34 34 47 51 53 56 60 83 85 82 83 80 78 78 79 94 Magqanchi-Quezon
[B] 66 64 80 85 40 42 53 48 58 64 71 86 88 87 88 90 90 92 94 84 79 Magqanchi-Patal Anew
[B] 66 65 80 85 39 41 52 47 60 65 72 86 90 89 90 91 93 93 94 85 81 94 Magqanchi-Yamot
[B] 63 61 76 82 37 39 49 47 58 63 70 87 89 87 87 88 86 89 91 84 79 92 93 Magqanchi-Patalkako
[B] 65 64 80 83 42 42 50 48 64 66 75 78 77 76 84 82 81 81 82 75 70 81 83 80 Abarlen-Pilien
[B] 53 53 65 68 33 34 42 44 56 57 61 77 79 77 79 78 79 76 77 83 76 77 79 77 70 Magqindi-Patalapalit
[B] 57 57 68 73 36 37 47 47 61 62 66 81 83 82 83 83 83 82 83 90 85 81 82 81 72 91 Magqindi-Balyew
[B] 55 53 66 70 35 36 45 45 59 58 64 78 81 79 80 78 8o 78 79 85 79 79 80 79 70 90 95 Magqindi-Lipitp
[B] 55 54 63 66 34 35 44 46 58 58 62 77 80 78 78 77 78 75 76 84 78 77 79 78 70 89 89 87 Magqindi-Malanak
[B] 55 54 67 70 37 38 45 45 60 59 65 75 79 78 77 79 83 77 79 82 78 78 80 77 71 89 89 89 95 Magqindi-Luwibao
[B] 58 57 70 73 37 39 45 45 61 62 66 77 81 80 79 80 85 81 80 84 77 80 82 80 73 88 89 87 86 96 Magqindi-Maagwage
[B] 54 53 64 67 33 36 46 49 55 56 61 76 76 75 75 75 74 73 74 84 77 74 76 75 69 86 87 83 84 82 82 Magqindi-Nabaklac
[B] 61 59 72 74 39 38 55 48 56 56 62 72 73 72 74 74 77 74 75 73 69 75 75 74 70 68 74 70 69 70 71 71 Ambala-Alibang
[B] 63 61 77 80 42 42 54 50 60 62 67 76 78 78 79 80 82 81 81 79 74 82 82 80 76 78 82 77 75 80 81 80 90 Ambala-Pili
[C] 59 57 69 70 38 38 54 47 56 55 57 67 67 66 68 68 70 67 67 67 64 67 67 65 66 64 69 65 64 67 69 67 86 85 Ambala-Malivecat
[B] 60 57 69 70 39 38 64 52 54 55 58 67 68 67 69 72 70 67 68 68 64 68 68 67 68 65 69 67 66 67 68 69 86 83 83 Ambala-Payangan
[B] 61 57 70 71 38 37 54 52 54 55 58 71 69 68 72 70 71 70 71 70 69 70 70 70 68 65 69 67 66 66 67 70 86 83 81 85 Ambala-Malabni
[B] 59 57 69 71 37 37 55 50 55 55 59 69 70 69 70 71 72 70 71 72 69 71 72 70 66 68 73 70 69 70 70 71 92 87 87 87 87 Ambala-Cabalan
[B] 56 55 63 63 40 37 63 49 51 51 55 61 61 61 61 64 63 61 63 60 57 64 64 61 61 54 56 55 56 55 56 56 67 65 68 73 65 70 Magbaken-Marang
[B] 53 51 59 60 38 35 61 47 49 49 53 58 59 58 59 60 60 69 60 58 55 62 61 59 58 55 59 57 58 57 59 58 70 69 72 75 68 73 99 Magbaken-Matanglaw

4 The Zambales Ayta Language Tree

The circled areas of Figure 2 indicate the proposed major groupings derived from the word list comparisons. The six languages defined are:

Ayta Language	Population[5]
1. Mag-anchi	4166
2. Aberlen	184
3. Mag-indi	2483
4. Ambala	1657
5. Mabeken	572
6. Abelen	2000

TABLE 2 - Major Ayta Language Groupings

It is anticipated that dialects from villages within a given language grouping will exhibit a fair amount of intelligibility; village dialects separated by these language boundaries should prove inherently unintelligible.

The tree derived from the cognate matrix is presented in Figure 3. The cognate percentages are listed in the appendix.

5 Summary

The Zambales Mountains contain · different Negrito languages, Abelen, Magqanchi, Abe , Magqindi, Ambala, and Mabeken. An assessment of the ility of the speakers of these languages to understand their neighbors and the trade languages is the goal of the second phase of the survey. Though one cannot predict intelligibility on the basis of cognate scores alone, the groupings presented should provide a starting point for further examination of language use patterns among these Zambales Ayta.

APPENDIX

AYTA LEXICOSTATISTIC MATRIX

Accuracy codes for each word list are given in brackets; see note 3 for an explanation of their meaning. The cognate percentage between two lists is read as the intersection between their names. For example, the Abelen-Dalayap and Magqanchi-Doray lists are 63 percent cognate.

[C] Sambal-Iba
[C] 81 Sambal-Santa Cruz
[B] 75 74 Sambal-Cabangan
[A] 76 74 96 Sambal-Botolan
[B] 44 44 45 46 Ilocano
[B] 41 44 44 44 47 Pangasinan
[B] 48 47 54 53 45 38 Tagalog
[C] 44 43 48 47 35 34 53 Pampanga
[C] 57 56 64 66 38 38 44 37 Abelen-Dalayap
[B] 58 57 66 68 38 40 44 38 93 Abelen-Tala
[C] 60 58 68 73 39 40 45 42 88 78 Abelen-Tangan
[B] 60 58 73 76 36 37 50 48 58 61 65 Magqanchi-Flora
[A] 60 58 75 78 40 37 50 49 63 62 67 83 Magqanchi-Buraq
[A] 60 59 75 78 40 38 49 47 63 63 67 80 98 Magqanchi-Inaaro
[B] 66 64 79 82 42 39 51 51 63 63 69 82 93 90 Magqanchi-Doray
[B] 65 64 82 86 40 39 51 47 65 66 72 86 88 88 88 88 Magqanchi-Kaiolkol
[B] 67 80 84 88 41 42 51 46 66 68 72 85 88 89 87 89 Magqanchi-Kaklingan
[A] 64 64 83 85 40 40 50 47 65 67 72 83 86 87 87 88 95 Magqanchi-San Martin
[A] 65 64 83 87 41 41 51 47 65 67 72 85 87 87 89 90 95 97 Magqanchi-San Isidro
[A] 58 58 88 73 36 36 48 50 80 81 84 88 91 88 87 84 83 81 83 Magqanchi-Sapangpak
[B] 55 63 55 69 34 34 47 51 53 56 80 83 85 82 83 80 78 78 79 94 Magqanchi-Quezon
[B] 66 64 80 85 40 42 53 48 58 64 71 86 88 87 88 90 90 92 94 84 79 Magqanchi-Patal Anaw
[B] 66 65 80 85 38 41 52 47 80 65 72 86 90 88 90 91 88 93 94 85 81 96 Magqanchi-Yaxat
[B] 63 61 76 82 37 38 49 47 58 63 70 87 88 87 87 88 88 88 91 84 79 92 93 Magqanchi-Patalkaco
[B] 65 64 80 83 42 42 50 48 64 66 75 78 77 76 84 82 81 81 82 75 70 81 83 80 Abarlan-Pilian
[B] 53 53 65 68 33 34 42 44 56 57 61 77 79 77 79 78 79 76 77 83 76 77 79 77 70 Magqindi-Patalapalit
[B] 57 57 68 73 36 37 47 47 61 62 66 81 83 82 83 83 83 82 83 90 85 81 82 81 72 91 Magqindi-Bahyaw
[B] 55 53 66 70 35 36 45 45 59 58 64 78 81 79 80 78 80 78 79 85 79 79 80 79 70 90 95 Magqindi-Lipilip
[B] 55 54 63 66 34 35 44 46 58 58 62 77 80 78 78 77 78 75 76 84 78 77 79 78 70 89 89 87 Magqindi-Malanak
[B] 55 54 67 70 37 38 45 45 60 59 65 75 78 78 77 79 83 77 79 82 76 78 80 77 71 89 89 89 85 Magqindi-Luzibao
[B] 58 57 70 73 37 38 45 45 61 62 66 77 81 80 79 80 85 81 80 84 77 80 82 80 73 88 88 87 86 96 Magqindi-Maagaga
[B] 54 53 64 67 33 36 46 49 55 56 61 76 76 75 75 75 74 73 74 84 77 74 76 75 69 86 87 83 84 82 82 Magqindi-Nabaklac
[B] 61 59 72 74 39 38 55 48 56 56 62 72 73 72 74 74 77 74 75 73 69 75 75 74 70 68 74 70 69 70 71 71 Aeta-Alitang
[B] 63 61 77 80 42 42 54 50 80 82 67 76 78 78 79 80 82 81 81 79 74 82 82 80 76 78 82 77 75 80 81 80 90 Aeta-Pili
[C] 59 57 69 70 38 38 54 47 56 55 57 67 67 66 68 68 70 67 67 67 64 67 67 65 66 64 69 65 64 67 69 67 86 85 Aeta-Maliwacat
[B] 60 57 69 70 39 38 64 52 54 55 58 67 68 67 69 72 70 67 68 68 64 68 68 67 88 85 89 87 66 67 68 69 86 83 83 Aeta-Poyanan
[B] 61 57 70 71 38 37 54 52 54 55 58 71 68 68 72 70 71 70 71 70 69 70 70 70 68 65 89 87 66 66 67 70 86 83 81 85 Aeta-Malabni
[B] 58 57 69 71 37 37 55 50 55 55 59 69 70 69 70 71 72 70 71 72 69 71 72 70 66 68 73 70 69 70 70 71 92 87 87 87 87 Aeta-Cabalan
[B] 56 55 63 63 40 37 63 49 51 51 55 61 61 61 61 64 63 61 63 80 57 64 64 61 61 54 56 55 56 55 56 56 67 65 68 73 65 70 Magbukun-Maranao
[B] 53 51 59 60 38 35 61 47 49 49 53 58 59 58 59 60 60 69 60 58 55 62 61 59 58 55 59 57 58 57 59 58 70 69 72 75 68 73 99 Magbukun-Manangloy

Notes

1. The word "cognate" in comparative linguistics normally refers to language forms which have descended from a single proto-form. The surveyor does not perform the rigorous analysis necessary to determine true cognates. Instead, he groups elicited forms into sets based on their similarities and his knowledge of the language family. Cognate, as used in survey and in this paper, should be understood as referring to these quickly determined "apparent" cognates.

2. These empirical figures are rules of thumb used by SIL surveyors in the Philippines. Simons (1983:67-100) explores several models which use lexicostatistics to predict intelligibility, and from which such figures might be derived.

3. Lawrence Reid developed the word list in 1966 by combining the relevant terms from the Swadesh 200 list with concepts central to Philippine life. The list has been used extensively by members of SIL in conducting language surveys, and is reproduced by Reid (1971) with data from some forty Philippine languages.

Most of the word lists were collected by Kurt and Margarette Storck, though a few were elicited or checked by the author. Word lists which had been obtained from previous surveys were rechecked. The accuracy codes in the appendix reflect the checking. A list was given the grade [C] when first obtained (either from archives or elicitation). A thorough recheck merited an upgrade to [B]. A grade of [A] indicates that the list was subjected to a second recheck. The compilation of multiple forms on the computer provided opportunity to recognize, and therefore recheck, suspicious items.

4. The WORDSURV program is available in the Summer Institute of Linguistics software library in the International Linguistic Center in Dallas, Texas. Several similar programs have been developed for use with small lap-top computers (Simons 1984), but these lack the full screen editing capabilities of WORDSURV, and can only handle a few word lists at a time. WORDSURV is only available for IBM-compatible computers and requires a 25-line display.

5. Population figures for the first five languages listed are from Storck 1986, as taken from government census. The figure for Abelen is an estimate taken from Wimbish 1984.

References

Busenitz, Bob. 1973. Sambal Survey. Ms. Summer Institute of Linguistics. Manila.

Casad, Eugene H. 1974. Dialect intelligibility testing. Norman, Oklahoma. Summer Institute of Linguistics.

Headland, Thomas N. 1975. Report of eastern Luzon language survey. Philippine Journal of Linguistics. Vol. 6, No. 2. Manila.

Kyle, Marc. 1977. Sambal Survey Report. Ms. Summer Institute of Linguistics. Manila.

Reid, Lawrence A. 1971. Philippine minor languages: word lists and phonologies. Oceanic linguistics special publication no. 8. University of Hawaii Press.

Savage, Dale. 1982. Pre-report memo -- Sambal / Ayta survey. Summer Institute of Linguistics. Manila.

Simons, Gary F. 1983. Language Variation and Limits to Communication. Summer Institute of Linguistics. Dallas.

-----. 1984. Word list analysis with a notebook computer. Occasional Publications in Academic Computing. No. 5. Summer Institute of Linguistics. Dallas.

Storck, Kurt and Margarette. 1986. A sociolinguistic survey report of the Ayta of the Western Cordillera of Luzon Island, Philippines. Ms. Summer Institute of Linguistics.

Walton, Charles. 1977. A Philippine language tree. Presented at the Austronesian Symposium. University of Hawaii. Honolulu.

Wimbish, John S. 1984. Abelen survey report. Ms. Summer Institute of Linguistics.